NBER WORKING PAPER SERIES

OPTIMAL ADAPTIVE CONTROL METHODS FOR STRUCTURALLY VARYING SYSTEMS

Alexander H. Sarris*

Michael Athans**

Working Paper No. 24

COMPUTER RESEARCH CENTER FOR ECONOMICS AND MANAGEMENT SCIENCE
National Bureau of Economic Research, Inc.
575 Technology Square
Cambridge, Massachusetts 02139

December 1973

*Massachusetts Institute of Technology and NBER Computer Research Center.
Research supported in part by National Science Foundation Grant GJ-1154X2 to the National Bureau of Economic Research, Inc.

**Massachusetts Institute of Technology. Research supported in part by NASA Grant NGL-22-009-124 to MIT.

Abstract

The problem of simultaneously identifying and controlling a time-varying, perfectly-observed linear system is posed. The parameters are assumed to obey a Markov structure and are estimated with a Kalman filter. The problem can be solved conceptually by dynamic programming, but even with a quadratic loss function the analytical computations cannot be carried out for more than one step because of the dual nature of the optimal control law. All approximations to the solution that have been proposed in the literature, and two approximations that are presented here for the first time, are analyzed. They are classified into dual and non-dual methods. Analytical comparison is untractable; hence Monte Carlo simulations are used. A set of experiments is presented in which five non-dual methods are compared. The numerical results indicate a possible ordering among these approximations.

Acknowledgement

The authors would like to thank Ms. Sophia A. Zalk for her excellent typing work.

Contents

Tables

Figures

1. INTRODUCTION

Economic science attempts to understand the economic behaviour of individual units like the household and the firm as well as their aggregates. There is huge diversity in the ways of people and firms, hence there is a lot of uncertainty inherent in any economic system. The difficulty of understanding economic behaviour is compounded by the fact that attitudes change, and technological innovations and political factors tend to always change the status quo. We live in a changing world and we must find ways to understand, describe, and deal with these changes.

To date most quantitative economic research has dealt with system models in which the structure is completely fixed and is not allowed to change. There has been a lot of work, under the name of econometrics, that has dealt with constant parameter estimation of econometric models. A very good indicator of the state of the art is the book by Theil (1971).

Recently there has been some research into the development of methods of describing and estimating changing parameters. The work of Rosenberg (1968), Cooley (1971) and Sarris (1973) are representative of the research to date.

This paper deals with policy in the presence of structural uncertainty as evidenced by parameter variations. There has been some research into the problem of policy formulation in the presence of constant but unknown system parameters. Prescott (1967) was the first economist to deal with such an "adaptive" problem. Since then McRai (1972), Poporic (1972), Rauser and Freebairn (1973), and Chow (1973) have also dealt with the same problem.

The problem of controlling a plant with unknown parameters is not new to engineers. Fel'dbaum (1960 a, b, 1961 a, b) was the first one to analyze the complexities of "learning while controlling," i.e. the dual nature of control. Since then there have been numerous books (Sworder (1966), Fel'dbaum (1966), Aoki (1967)) and papers (see ref. 16 for an extensive bibliography) dealing with policy in the presence of uncertain parameters. However, there have been very few papers, addressing themselves specifically to the problem of controlling a system whose parameters are varying in a random fashion. Exceptions are the papers by Wieslander and Wittenmark (1971) and Wouters (1972), in which some numerical results were given. The papers by Bar-Shalom and Sivan (1969), Tse and Athans (1972), Tse et. al. (1973 a,b) also treated time varying parameters although the numerical results reported were for systems with constant parameters.

In this paper we attempt to unify most of the methods available for controlling systems with parameter adaptation. To this end we shall consider only systems with perfect state information. We shall extend the methods that have been developed for the constant parameter case, to include the varying parameter case. We shall also propose some new methods. In section 2 we present the problem to be tackled. Section 3 analyzes the estimation technique for the time varying parameters. Section 4 presents the general method of solution and indicates the difficulties of applying it to our problem. In section 5, we present the ideal case of known parameters and one control technique based on it. In section 6 we present four non-dual control methods and try to indicate their shortcomings. In section 7 we present three dual methods, one of which is presented here for the first time. Section 8 presents some Monte Carlo comparisons of the non-dual methods, and in section 9 we summarize the results and indicate directions for further research.

2. STATEMENT OF THE PROBLEM

Our purpose is to analyze and compare various methods so we shall try to keep the complexity of the systems to be analyzed, minimal. Generalizations of the methods to more complicated problems are straight forward in most cases. We shall confine ourselves to discrete time linear systems described by the following equations.

$$x_{t+1} = A_t x_t + B_t u_t + C_t z_t + \varepsilon_t \qquad (1)$$
$$y_t = H_t x_t + v_t \qquad (2)$$

where

x_t - is the unobservable state vector at time t

u_t - is a vector of policy or control variables at time t

z_t - is a vector of exogenous variables

y_t - is the vector of state measurements at time t

ε_t, v_t - are vectors of system and measurement noises respectively.

The model as stated in (1) and (2) is general enough to include many engineering models of interest and also reduced form econometric models. However, it is still too general for our purposes. Therefore, we shall consider the following model composed of the most elementary building blocks.

$$y_{t+1} = a_t y_t + b_t u_t + \varepsilon_t \qquad (3)$$

where

y_t - is the perfectly observed scalar state

u_t - is a scalar control

ε_t - is scalar system noise.

The model in (3) is a special case of almost every model that has been dealt with in the literature. Hence we can compare many methods at this level.

The state y_t will be measured exactly. Let us denote by y^t, u^t the following quantities

$$y^t \equiv \{y_o, y_1, \ldots .. y_t\} \qquad (4)$$

$$u^t \equiv \{u_o, u_1, \ldots .. u_t\} \qquad (5)$$

The controls $\{u_t\}$ will be restricted to the following form.

$$u_t = \gamma_t(y^t, u^{t-1}) \qquad (6)$$

where γ_t is a function to be chosen. Let Y_t denote the set in which the state at time t is restricted to lie, and V_t the set of allowable u_t's. Then γ_t is a function from $Y_o \times Y_1 \ldots \times Y_t \; V_o \times V_1 \ldots V_{t-1}$ into V_t. For the purposes of this paper $Y_i = V_i = R$ for all i.

At time zero we shall assume that the following quantities are known;

$$y_o, \; p(\varepsilon_o, \varepsilon_1, \ldots, \varepsilon_N) = \prod_{i=0}^{N} N(0, \sigma_\varepsilon^{\;2})^*.$$

$$p(a_o, b_o) = N\left[\begin{pmatrix} \bar{a}_o \\ \hline \bar{b}_o \end{pmatrix}, M_o\right]$$

The objective is to choose the functions $\gamma_o, \gamma_1, \ldots, \gamma_{N-1}$ such that the following cost criterion is minimized.

$$V(y_o) = E\left\{\sum_{i=0}^{N=1} (y_{i+1}^2 + ru_i^2)\right\} \qquad (6a)$$

* p(.) denotes a probability density and N(a,b) denotes a normal density with mean a and variance b.

Notice that the problem is still not completely formulated because we do not know how a_t and b_t are going to vary. We shall impose the following a-priori probabilistic structure on the parameters.

$$\begin{bmatrix} a_t \\ b_t \end{bmatrix} = \begin{bmatrix} a_{t-1} \\ b_{t-1} \end{bmatrix} + \begin{bmatrix} \theta_{t-1} \\ \eta_{t-1} \end{bmatrix} \quad (7)$$

or if we denote by $p_t \equiv [a_t \ b_t]'$ * and $w_t = [\theta_t, \eta_t]'$

$$p_t = p_{t-1} + w_{t-1} \quad (8)$$

This is the structure proposed by Rosenberg (1968) and Sarris (1973).

In order for the problem to be completely specified the joint probability density of $W_o, W_1, \ldots W_N$ must be given. Since we do not know a-priori how the parameters vary it is not trivial to specify this quantity. For the purposes of this paper we shall make the following assumption

$$p(W_o, W_1, \ldots W_N) = \sum_{i=1}^{N} N(0,R) \quad (9)$$

where

$$R = \begin{bmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix} \quad (10)$$

The choice of appropriate R will not be discussed in this paper. It is discussed somewhat by Sarris (1973).

The problem can now be stated in full.

* (') denotes transposition

Find the optimum $V^*(y_o)$ where

$$V^*(y_o) = \min_{\gamma_0,\gamma_1,\ldots\gamma_{N-1}} E \left\{ \sum_{i=0}^{N-1} (y_{i+1}^2 + ru_i^2) \right\} \qquad (11)$$

subject to the stochastic constraints,

$$y_{t+1} = a_t y_t + b_t u_t + \varepsilon_t \qquad (12)$$

$$\begin{bmatrix} a_t \\ b_t \end{bmatrix} \equiv p_t = p_{t-1} + w_{t-1} \qquad (13)$$

where $\{\varepsilon_t\}$ and $\{w_t\}$ are series of white normal random variables with the properties

$$p(\varepsilon_t) = N(0,\sigma_\varepsilon^2) \qquad (14)$$

$$p(w_t) = N(0, R) \qquad (15)$$

$$p(\varepsilon_i,w_j) = p(\varepsilon_i)p(w_j) \qquad (16)$$

and the system initial conditions are,

$$y_o \text{ - known} \qquad (17)$$

$$p(p_o) = N(\bar{p}_o, M_o) \qquad (17a)$$

In the sequel we will abuse the notation a little by writing $u_o, u_1, \ldots, u_{N-1}$ in place of $\gamma_o,\gamma_1,\ldots\ldots\gamma_{N-1}$ in (11). This will be done for the reader's convenience.

3. BAYESIAN ESTIMATION OF THE VARYING PARAMETERS

As will be seen soon, the solution of the problem stated in section 2 will require the knowledge of the joint conditional distribution of the parameters a_t and b_t, conditioned on the data up to the time t. In this section we shall examine a way of obtaining this distribution, which we shall denote by $p(p_t/y^t,u^{t-1})$.

The distribution at time zero is normal as seen in (17a). Assume that the conditional distribution $p(p_{t-1}/y^t, u^{t-1})$ is normal with mean denoted by $p_{t-1/t-1}$, and symmetric covariance matrix denoted by $M_{t-1/t-1}$. The relevant equations for the next stage are

$$p_t = p_{t-1} + w_{t-1} \tag{18}$$

$$y_{t+1} = z_t p_t + \varepsilon_t \tag{19}$$

where we have denoted

$$z_t \equiv [y_t \ u_t] \tag{20}$$

We can use standard Bayesian analysis to find the density $p(p_t/y^{t+1},u^t)$

$$p(p_t/y^{t+1},u^t) = \frac{p(y_{t+1}/p_t,y^t,u^t)p(p_t/y^t,u^t)}{\int p(y_{t+1}/p_t,y^t,u^t)p(p_t/y^t,u^t)dp_t} \tag{21}$$

From (18) we see that the density $p(p_t/y^t,u^t) = p(p_t/y^t,u^{t-1})$ is normal with mean equal to $p_{t/t-1} = p_{t-1/t-1}$ and covariance matrix

$$M_{t/t-1} = M_{t-1/t-1} + R \tag{22}$$

$p(y_{t+1}/p_t,y^t,u^t)$ is also normal from (19). The density in (21) is therefore normal. Its mean and covariance matrix, after some calculations, are given by the following formulas.

$$P_{t/t} = M_{t/t}\left[M_{t/t-1}^{-1} P_{t/t-1} + \frac{1}{\sigma_\varepsilon^2} z_t' y_{t+1}\right] \quad (23)$$

$$M_{t/t}^{-1} = M_{t/t-1}^{-1} + \frac{1}{\sigma_\varepsilon^2} z_t' z_t \quad (24)$$

The following matrix inversion lemma will help us render (23) and (24) identical to the standard Kalman filter equations.

Lemma 1. (Matrix Inversion Lemma). If

$$S = [M^{-1} + AR^{-1}B]^{-1} \quad \text{then}$$

$$S = M - MA[R + BMA]^{-1} BM.$$

Proof. The proof is by direct computation and is omitted.

With the help of this lemma (24) can be rewritten.

$$M_{t/t} = M_{t/t-1} - M_{t/t-1} z_t' [\sigma_\varepsilon^2 + z_t M_{t/t-1} z_t']^{-1} z_t M_{t/t-1} \quad (25)$$

This along with (22) are the well known updating equations for the covariances of the Kalman filter adapted to our problem. We notice that since $M_{t/t-1}$ is symmetric then $M_{t/t}$ is also symmetric. We now substitute in (23) the expression for $M_{t/t}$ found in (25). We obtain after some manipulation

$$P_{t/t} = P_{t/t-1} + M_{t/t-1} z_t' (\sigma_\varepsilon^2 + z_t M_{t/t-1} z'_t)^{-1} (y_{t+1} - z_t P_{t/t-1}) \quad (26)$$

Which is the standard Kalman updating formula. Equations (23) and (24) will be useful later.

4. SOLUTION VIA DYNAMIC PROGRAMMING

The problem that was stated in section 2 can in principle be solved via dynamic programming. We state now the form that the stochastic dynamic programming equations take. We can write:

$$V^*(y_o) = \min_{u_o,\, u_1, \ldots, u_{N-1}} E' \{E[\sum_{i=0}^{N-1} (y_{i+1}^2 + ru_i^2)/y^{N-1}, u^{N-2}]\} \quad (27)$$

We shall now state a theorem, which can be found in Åström (1970, ch. 8), that will be crucial.

<u>Theorem 1.</u> Let $E\,[./y]$ denote the conditional mean given y. Assume that the function $f(y,u) = E[l(x,y,u)/y]$ has a unique minimum with respect to $u\varepsilon V$ for all $y\varepsilon Y$. Let $u^o(y)$ denote the value of u for which the minimum is achieved. Then

$$\min_{u(y)} El(x,y,u) = El(x,y,u^o(y)) = E'_y \{\min_u E\,[l(x,y,u)/y]\}$$

Using this theorem and noticing that $E[\sum_{i=0}^{N-1} (y_{i+1}^2 + ru_i^2)/y^{N-1}, u^{N-2}]$ is quadratic with respect to u_{N-1}, therefore having a unique minimum we can write

$$V^*(y_o) = \min_{u_o, u_1, \ldots, u_{N-2}} E' \{\min_{u_{N-2}} E[\sum_{i=0}^{N-1} (y_{i=1}^2 + ru_i^2)/y^{N-1}, u^{N-2}]\} \quad (28)$$

Now we invoke the principle of optimality, and noticing that the first N-1 terms in the summation of (28) do not involve u_{N-1} , we write:

$$V^*(y_o) = \min_{u_o,u_1,\ldots u_{N-2}} E'\{\sum_{i=0}^{N-2}(y_{i+1}^2 + ru_i^2) + \min_{u_{N-1}} E[y_N^2 + ru^2_{N-1}/y^{N-1},u^{N-2}]\}$$

$$\equiv \min_{u_o,u_1,\ldots u_{N-2}} E'\{\sum_{i=t}^{N-2}(y_{i+1}^2 + ru_i^2) + V^*(y^{N-1})\} \qquad (29)$$

where we have denoted:

$$V^*(y^t) = \min_{u_t,u_{t+1},\ldots u_{N-1}} E'\{\sum_{i=t}^{N-1}(y_{i+1}^2 + ru_i^2)/y^t,u^{t-1}\}$$

By the reasoning used above it is quite straightforward now to prove the following recursive relation:

$$V^*(y^t) = \min_{u_t} E'\{y_{t+1}^2 + ru_t^2 + V^*(y^{t+1})/y^t,u^{t-1}\} \qquad (30)$$

Equation (30) is the well known recursive relation of stochastic dynamic programming. If we can solve it then our problem will be solved.

At time N-1 (30) becomes:

$$V^*(y^{N-1}) = \min_{u_{N-1}} E\{y_N^2 + ru_{N-1}^2/y^{N-1},u^{N-2}\} =$$

$$\min_{u_{N-1}} E\{(a_{N-1}^2y_{N-1}^2 + b_{N-1}^2u_{N-1}^2 + \varepsilon_{N-1}^2 + 2a_{N-1}b_{N-1}u_{N-1}y_{N-1} +$$

$$2a_{N-1}y_{N-1}\varepsilon_{N-1} + 2b_{N-1}u_{N-1}\varepsilon_{N-1}) + ru_{N-1}^2/y^{N-1},u^{N-2}\} =$$

$$= \min_{u_{N-1}} \Big[y_{N-1}^2 E(a_{N-1}^2/y^{N-1},u^{N-2}) + u_{N-1}^2 E(b_{N-1}^2/y^{N-1},u^{N-2}) +$$

$$+ \sigma_\varepsilon^2 + 2u_{N-1}y_{N-1}E(a_{N-1}b_{N-1}/y^{N-1},u^{N-2}) + ru_{N-1}^2\Big] \qquad (31)$$

The minimum of the above expression is easy to find since the quantity inside the brackets is a quadratic in u_{N-1}.

$$u_{N-1}^* = -\left[r + E(b_{N-1}^2/y^{N-1},u^{N-2})\right]^{-1} E(a_{N-1}b_{N-1}/y^{N-1},u^{N-2})y_{N-1} \qquad (32)$$

$$V^*(y^{N-1}) = K_{N-1}\, y_{N-1}^2 + L_{N-1} \qquad (33)$$

where

$$K_{N-1} = E(a_{N-1}^2/y^{N-1},u^{N-2}) - \left[r + E(b_{N-1}^2/y^{N-1},u^{N-2})\right]^{-1} E(a_{N-1}b_{N-1}/y^{N-1},u^{N-2})^2 \qquad (34)$$

$$L_{N-1} = \sigma_\varepsilon^2 \qquad (35)$$

Equation (33) might look like a quadratic in y_{N-1} but a quick look at (34) will convince the reader that K_{N-1} is a quite complicated function of y_{N-1} (c.f. equations (25)÷(26)). It thus becomes impossible to carry the backward induction any further than already. done.

It is our purpose in this paper to examine and compare suboptimal techniques to solve the problem posed in section 2. This will be done in the next few sections.

5. OPTIMAL CONTROL WITH PERFECTLY KNOWN PARAMETERS

In this section we shall assume that the parameters a_t, b_t are known with certainty during the whole interval $[0,N]$. Equation (30) at time N-1 becomes:

$$V^*(y^{N-1}) = \min_{u_{N-1}} E' \{y_N^2 + ru_{N-1}^2 / y^{N-1}, u^{N-2}\} =$$

$$\min_{u_{N-1}} E' \{ a_{N-1}^2 y_{N-1}^2 + b_{N-1}^2 u_{N-1}^2 + 2a_{N-1}b_{N-1}u_{N-1}y_{N-1} + \varepsilon_{N-1}^2 +$$

$$2a_{N-1}y_{N-1}\varepsilon_{N-1} + 2b_{N-1}u_{N-1}\varepsilon_{N-1} + ru_{N-1}^2 / y^{N-1}, u^{N-2} \} =$$

$$\min_{u_{N-1}} \left[a_{N-1}^2 y_{N-1}^2 + b_{N-1}^2 u_{N-1}^2 + 2a_{N-1}b_{N-1}u_{N-1}y_{N-1} + \sigma_\varepsilon^2 + ru_{N-1}^2 \right] \quad (36)$$

The above equation is a quadratic in u_{N-1} so its minimum is easily found.

$$u_{N-1}^* = - \left[r + b_{N-1}^2\right]^{-1} a_{N-1}b_{N-1}y_{N-1} \quad (37)$$

$$V^*(y^{N-1}) = H_{N-1}y_{N-1}^2 + F_{N-1} \quad (38)$$

where

$$H_{N-1} = a_{N-1}^2 - \left[r + b_{N-1}^2\right]^{-1} a_{N-1}^2 b_{N-1}^2 \quad (39)$$

$$F_{N-1} = \sigma_\varepsilon^2$$

Let $V^*(y^{j+1}) = H_{j+1}y^2_{j+1} + F_{j+1}$. Then at time t=j the dynamic programming recursion becomes

$$V^*(y^j) = \min_{u_j} E\left[y_{j+1}^2 + ru_j^2 + H_{j+1}y_{j+1}^2 / y^j, u^{j-1}\right] =$$

$$\min_{u_j} \left[(1 + H_{j+1})(a_j^2 y_j^2 + b_j^2 u_j^2 + 2a_j b_j u_j y_j + \sigma_\varepsilon^2) + ru_j^2\right] \quad (40)$$

The minimum of the above equation is again easily found:

$$u_j^* = - \left[r + (1 + H_{j+1})b_j^2\right]^{-1} (1 + H_{J+1})a_j b_j y_j \quad (41)$$

$$V^*(y^j) = H_j y_j^2 + F_j \quad (42)$$

where

$$H_j = (1 + H_{j+1})a_j^2 - \left[r + (1 + H_{j+1})b_j^2\right]^{-1}(1 + H_{j+1})^2 a_j^2 b_j^2 \quad (43)$$

$$F_j = F_{j+1} + (1 + H_{j+1})\sigma_\varepsilon^2 \quad (44)$$

The equations (41)-(44) along with the initial conditions $F_N=0$ and $H_N=0$ are the solution to the problem.

A suboptimal technique of solving the original problem is based on (41)-(44) and is usually referred to in the literature by the name of certainty equivalence or enforced separation (from here on abbreviated as CE). It is the following,

a) At time k we are given the data y^k and u^{k-1} hence the following quantities can be computed via the results of section 3:

$$P_{k/k-1} = \left[a_{k/k-1}, b_{k/k-1}\right]' \text{ and } M_{k/k-1}$$

b) Equation (43) is solved backward from time N until time k+1 with the following conventions:

1) $\left.\begin{aligned} a_j &= a_{k/k-1} \\ b_j &= b_{k/k-1} \end{aligned}\right\}$ for all $k+1 \leq j \leq N-1$

2) $H_N = 0$

Denote the solution by H_{k+1}^{CE}

c) The control at time k is found by the following equation:

$$u_k^{CE} = -\left[r + (1 + H_{k+1}^{CE})b_{k/k-1}^2\right]^{-1}(1 + H_{k+1}^{CE})a_{k/k-1}b_{k/k-1}y_k \qquad (45)$$

This suboptimal technique is usually the one against which most people compare their suboptimal methods. It is one of the simplest and fastest suboptimal techniques and therefore it is attractive. It will be compared with other suboptimal methods at a later section. It is interesting to see that if the parameters are known exactly CE reduces to the true control law (41).

6. NON-DUAL SUBOPTIMAL METHODS.

In this section we shall examine various suboptimal techniques that have been suggested in the literature. All these techniques will be non-dual, in the sense that they calculate the control law at time k under the assumption that there will be no further measurements after time k.

There are three main elements of a dual control. The first which can be called the controlling element has to do with the effect of the control on the criterion function and is the element that characterizes all optimum controls, dual or not. The second characteristic is a learning one, namely the information that is accumulated over past controlling stages is utilized to improve the present knowledge of the system. In section 3 we analyzed the way that optimal learning will be achieved in our problem. The third element, which we shall term the dual effect, has to do with the experimental nature of the control. Choices of present controls affect the future probability densities of the unknown parameters. Hence a dual control can affect not only the present but also the future learning of the system. It will be this element that will be missing from the suboptimal methods presented in this section. In all subsequent methods, learning will occur via the method described in section 3.

6.1 Wouters' Minimum Variance Control.

This method was proposed by Wouters (1972). It is quite simple. The logic is the following. Suppose that the objective is to minimize the index;

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} y^2(k) \qquad (46)$$

then the control suggested by Wouters (to be denoted by the letter W) is

$$u_k^W = - \frac{a_{k/k-1}}{b_{k/k-1}} y_k \qquad (47)$$

Notice that (46) is quite different than our objective (6). It does not, for example, include penalty for the control. Wouters used this technique to control systems with time varying parameters. He showed via Monte Carlo experiments that the method is better than no control at all.

6.2 Wieslander's and Wittenmark's Control.

This method (hereby denoted as WW) was proposed by Wieslander and Wittenmark (1971). Their idea is the following. Since the recursive equation (30) cannot be solved analytically for more than one step, assume that the next step is the final one. The index to be minimized in their paper was $Ey^2(t)$. The control that they derived is the same one as in (32) with r = 0

$$u_k^{WW} = -\left[E(b_k^2/y^k,(u^{WW})^{k-1})\right]^{-1} E(a_k b_k/y^k,(u^{WW})^{k-1})y_k \qquad (48)$$

In the experiments that they did they compared this control law to no control at all, and it performed better. Since it is not obvious that any control will perform better than no control, their method deserves some attention. This as well as the previous method ignores penalty in the control. However, in this case it is quite easy to introduce control penalty. In fact the modified control law (to be denoted by WM) is identical to the one in (32).

$$u_k^{WM} = -\left[r + E(b_k^2/y^k,(u^{WM})^{k-1})\right]^{-1} E(a_k b_k/y^k,(u^{WM})^{k-1})y_k \qquad (49)$$

It is interesting to notice that none of the previous three methods reduce to the true control law, derived in section 5 (equation (41)), when the parameters are known exactly. We now examine a method that has this desirable property.

6.3 Sequential Stochastic Control.

The logic for this method is that at time k all future information is neglected. However, it is recognized that the parameters will be changing. The assumption then is that the distribution of the future values of the parameters will not be affected by the future measurements. This assumption is similar to the one that assumes the future parameters to be random drawings from a distribution which depends only on information up to time k. The difference here is that the distribution is different at every point in time. This method has been mentioned by Yoshida and Nakamura (1973), but they have not analyzed it carefully. We now derive it in detail (the method will be abbreviated by S1).

Assume that we are at time k and we have observed y^k, u^{k-1}. Hence we have computed $p_{k/k-1}$ and $M_{k/k-1}$ with the help of the Bayesian formulas developed in section 3. The problem now is the following. Choose $u_k, u_{k+1}, \ldots u_{N-1}$ so as to minimize

$$V(y^k) = E'\{ \sum_{j=k}^{N-1} (y_{j+1}^2 + ru_j^2)/y^k, u^{k-1} \} \quad (50)$$

subject to

$$\left.\begin{aligned} y_{j+1} &= z_j p_j + \varepsilon_j \\ p_j &= p_{j-1} + w_{j-1} \end{aligned}\right\} \quad j \geq k \quad (51)$$

The assumption that we are making can now be stated precisely. The vector p_j of parameters at time $j \geq k$ will be assumed to be a random drawing from a Gaussian density with mean

$$p_{j/k-1} = p_{j-1/k-1} = \cdots = p_{k/k-1} \quad (52)$$

and covariance matrix

$$M_{j/k-1} = M_{j-1/k-1} + R = \ldots = M_{k/k-1} + (j-k)R \quad (53)$$

Thus we approximate $p(p_j/y^j,u^{j-1})$ by $p(p_j/y^k,u^{k-1})$. The dynamic programming recursion now can be analyzed. At the final time (30) becomes

$$V^*(y^{N-1}) = \min_{u_{N-1}} E' \{y_N^2 + ru_{N-1}^2 \;/\; y^{N-1},u^{N-1}\} =$$

$$\min_{u_{N-1}} E' \{ a_{N-1}^2 y_{N-1}^2 + b_{N-1}^2 u_{N-1}^2 + \varepsilon_{N-1}^2 + 2a_{N-1}b_{N-1}u_{N-1}y_{N-1} +$$

$$2a_{N-1}y_{N-1}\varepsilon_{N-1} + 2b_{N-1}u_{N-1}\varepsilon_{N-1} + ru_{N-1}^2/y^{N-1},u^{N-2} \} =$$

$$\simeq \min_{u_{N-1}} \Big[y_{N-1}^2 E(a_{N-1}^2/y^k,u^{k-1}) + u_{N-1}^2 E(b_{N-1}^2/y^k,u^{k-1}) + \sigma_\varepsilon^2 +$$

$$2u_{N-1}y_{N-1}E(a_{N-1}b_{N-1}/y^k,u^{k-1}) + ru^2_{N-1}\Big] \qquad (54)$$

Let us now decompose the matrix $M_{j/k-1}$ as follows:

$$M_{j/k-1} = \begin{bmatrix} M^a_{j/k-1} & M^{ab}_{j/k-1} \\ M^{ab}_{j/k-1} & M^b_{j/k-1} \end{bmatrix} \qquad (55)$$

Referring to (10) and (55), (54) reduces to

$$V^*(y^{N-1}) = \min_{u_{N-1}} \Big[y^2_{N-1}(a^2_{N-1/k-1} + M^a_{N-1/k-1}) + u^2_{N-1}\Big[r +$$

$$(b^2_{N-1/k-1} + M^b_{N-1/k-1})\Big] + \sigma_\varepsilon^2 + 2u_{N-1}y_{N-1}(a_{N-1/k-1}b_{N-1/k-1} + M^{ab}_{N-1/k-1})\Big] \qquad (56)$$

The control minimizing the above expression is

$$u^*_{N-1} = -\left[r + (b^2_{N-1/k-1} \; M^b_{N-1/k-1})\right]^{-1} (a_{N-1/k-1} b_{N-1/k-1} + M^{ab}_{N-1/k-1}) y_{N-1} \quad (57)$$

$$V^*(y^{N-1}) = H_{N-1} y^2_{N-1} + F_{N-1} \quad (58)$$

where

$$H_{N-1} = (a^2_{N-1/k-1} + M^a_{N-1/k-1}) - \left[r + (b^2_{N-1/k-1} + M^b_{N-1/k-1})\right]^{-1} \cdot (a_{N-1/k-1} b_{N-1/k-1} + M^{ab}_{N-1/k-1})^2 \quad (59)$$

$$F_{N-1} = \sigma^2_\varepsilon \quad (60)$$

If we now assume that

$$V^*(y^{j+1}) = H_{j+1} y^2_{j+1} + F_{j+1} \quad (61)$$

then by an analysis identical to that of section 5 we derive the following:

$$u^*_j = -\left[r + (1 + H_{j+1})(b^2_{j/k-1} + M^b_{j/k-1})\right]^{-1} (1 + H_{j+1}) \cdot (a_{j/k-1} b_{j/k-1} + M^{ab}_{j/k-1}) \, y_j \quad (62)$$

$$V^*(y^j) = H_j y^2_j + F_j \quad (63)$$

where

$$H_j = (1 + H_{j+1})(a^2_{j/k-1} + M^a_{j/k-1}) - \left[r + (1 + H_{j+1})(b^2_{j/k-1} + M^b_{j/k-1})\right]^{-1} (1 + H_{j+1})^2 (a_{j/k-1} b_{j/k-1} + M^{ab}_{j/k-1})^2$$

$$F_j = F_{j+1} + (1 + H_{j+1}) \, \sigma^2_\varepsilon \quad (65)$$

$$H_N = F_N = 0 \quad (66)$$

The optimal control at time k is chosen as follows:

$$u_k^{S1} = u_k^* \qquad (67)$$

where u_k^* is derived recursively as above. After this control is applied y_{k+1} is observed and the cycle is repeated to choose u_{k+1} and so on until time N-1. It is interesting to note that when the parameters are known exactly the control derived by this method is reduced to the true optimal control described in section 5. When R = 0 or equivalently when we assume that the parameters are constant, then S1 reduces to a method that has been analyzed among others by Aoki (1967), Bar-Shalom and Sivan (1969), and Prescott (1967).

6.4 Open Loop Feedback Optimal (OLFO) Control.

This method has been analyzed by Tse and Athans (1972) and Ku and Athans (1973). The assumption under which the control at time k is found is that the sequence $u_k, u_{k+1}, \ldots u_{N-1}$ will not depend on any future data and hence can be found at time k by solving an open loop control problem. Let us make this assumption more precise. The problem to be solved at time k is the following.

$$V^*(y^k) = \min_{u_k, u_{k+1}, \ldots u_{N-1}} \{ E[\sum_{j=k}^{N-1} y_{j+1}^2 / y^k, u^{k-1}] + r \sum_{j=k}^{N-1} u_j^2 \} \qquad (68)$$

subject to

$$\left.\begin{array}{l} y_{j+1} = z_j p_j + \varepsilon_j \\ p_j = p_{j-1} + w_{j-1} \end{array}\right\} \quad j \geq k \qquad (69)$$

Notice that the expectation in (68) does not include the control terms. This is because they are to be chosen in an open loop fashion. The solution to this problem is quite complicated. We shall present here an outline of it and we shall mention the simplifications that were employed by Tse and Athans, and Ku and Athans.

The problem in (68) and (69) can be solved via deterministic dynamic programming as follows. Denote by $V^*(y^j)$ the quantity

$$V^*(y^j) = \min_{u_j,u_{j+1},\ldots,u_{N-1}} E\{\sum_{i=j}^{N-1} (y_{i+1}^2 + ru_i^2)/y^k, u^{k-1}\}$$

$$\equiv \min_{u_j,u_{j+1},\ldots,u_{N-1}} E\{\sum_{i=j}^{N-1} (y_{i+1}^2 + ru_i^2)/k-1\} \qquad (70)$$

Then the dynamic programming recursion is

$$V^*(y^j) = \min_{u_j} E\{y_{j+1}^2 + ru_j^2 + V^*(y^{j+1}/k-1\} \qquad (71)$$

Notice that since $E(./k-1)$ is known at time k, (71) is a deterministic dynamic programming recursion.

At the final step we obtain

$$V^*(y^{N-1}) = \min_{u_{N-1}} E(y_N^2 + ru_{N-1}^2/k-1) =$$

$$= \min_{u_{N-1}} E(a_{N-1}^2 y_{N-1}^2 + b_{N-1}^2 u_{N-1}^2 + 2a_{N-1}b_{N-1}u_{N-1}y_{N-1} + 2a_{N-1}y_{N-1}\varepsilon_{N-1} +$$

$$2b_{N-1}u_{N-1}\varepsilon_{N-1} + ru_{N-1}^2/k-1) =$$

$$\min_{u_{N-1}}\{ E(a_{N-1}^2 y_{N-1}^2/k-1) + u_{N-1}^2 \left[r + E(b_{N-1}^2/k-1)\right] +$$

$$2u_{N-1}E(a_{N-1}b_{N-1}y_{N-1}/k-1) + \sigma_\varepsilon^2 \} \qquad (72)$$

The optimal OLFO u_{N-1} is

$$u_{N-1}^* = -D_{N-1}^{-1} f_{N-1} \qquad (73)$$

where

$$D_{N-1} = r + E(b_{N-1}^2/k-1) \qquad (74)$$

$$f_{N-1} = E(a_{N-1}b_{N-1}y_{N-1}/k) \qquad (75)$$

$$V^*(j^{N-1}) = E(a^2_{N-1}y_{N-1}^2/k-1) - D^{-1}_{N-1} f_{N-1}^2 + \sigma_\varepsilon^2 \qquad (76)$$

Notice an interesting phenomenon. Since in the state equations (69) a, b, and y are coupled in a nonlinear manner one cannot separate $E(a_{N-1}^2 y_{N-1}^2/k-1)$ for example into $E(a_{N-1}^2/k-1)E(y_{N-1}^2/k-1)$. Hence no interesting cancellations will occur in the steps prior to the last. To illustrate this point we will show without proof (which is straightforward) the OLFO control and the cost at time N-2.

$$u_{N-2}^* = -D^{-1}_{N-2}f_{N-2} \qquad (77)$$

$$D_{N-2} = r + E(b_{N-2}^2 + a_{N-1}^2 b_{N-2}^2/k-1) - D^{-1}_{N-1}E^2(a_{N-1}b_{N-1}b_{N-2}/k-1) \qquad (78)$$

$$f_{N-2} = E(a_{N-2}b_{N-2}y_{N-2} + a_{N-1}^2 a_{N-2}b_{N-2}y_{N-2}/k-1) - D_{N-1}^{-1}E(a_{N-1}b_{N-1}a_{N-2}y_{N-2}/k-1)\cdot$$

$$\cdot E(a_{N-1}b_{N-1}b_{N-2}/k-1) \qquad (79)$$

$$V^*(y^{N-2}) = E(a_{N-2}^2 y_{N-2}^2 + a_{N-1}^2 a_{N-2}^2 y_{N-2}^2 / k-1) - D_{N-1}^{-1} E^2(a_{N-1} b_{N-1} a_{N-2} y_{N-2} / k-1) - D_{N-2}^{-1} f_{N-2}^2 \quad (80)$$

Thus we can see that the exact solution for the OLFO control at time k becomes increasingly laborious as we proceed in the backwards induction. The problem arises because we have assumed that a_j as well as b_j are random, and this introduces the nonlinearity in (69). Tse and Athans (1972) assumed that only b_j is random while a_j is not. In such a case

$$\begin{aligned} E(y_{j+1}/k-1) &= a_j E(y_j/k-1) + u_j E(b_j/k) \\ E(b_{j+1}/k-1) &= E(b_j/k-1) \end{aligned} \quad (81)$$

and therefore the conditional expectations evolve linearly, making the backwards induction of reproducible form from step to step. Ku and Athans (1973) on the other hand have used the approximation

$$E(y_{j+1}/k-1) = E(a_j/k-1)E(y_j/k-1) + u_j E(b_j/k-1) \quad (82)$$

Their extensive Monte Carlo results showed that OLFO in conjunction with (82) performed slightly better than CE (or enforced separation, as they called CE), for stable systems, but considerably worse than CE for unstable ones.

7. DUAL SUBOPTIMAL METHODS.

Dual methods assume explicitly that the choice of the present control will affect the future probability densities of the parameters. Hence the control is inevitably a nonlinear function of the present state and in most cases quite a complicated one too. We shall analyze three quite different dual methods, the last one appearing here for the first time.

7.1 One-Measurement-Optimal Feedback Control.

This measurement was developed by Curry (1969-1970), and has been recently used by Tse et.al. (1973), Tse and Bar-Shalom (1973), Rausser and Freebairn (1973), and further analyzed by Early and Early (1973). The idea is the following.

Suppose we chose $u_k = \bar{u}_k$. Then we could find the covariance of p_k given $\{y^k, u^{k-1}, \bar{u}_k\}$ via (25). We could also assert that the average value of y_{k+1} would be

$$\bar{y}_{k+1} = y_{k+1/k-1} = a_{k/k-1} y_k + b_{k/k-1} \bar{u}_k \tag{83}$$

We could then consider the problem

$$V(y^k, \bar{y}_{k+1}) = \min_{u_{k+1}, \ldots, u_{N-1}} E' \{ \sum_{i=k+1}^{N-1} (y_{i+1}^2 + r u_i^2) / y^k, \bar{y}_{k+1}, u^{k-1}, \bar{u}_k \}$$

with initial conditions

$$\bar{y}_{k+1} = y_{k+1/k-1} \tag{84}$$

$$P_{k/k} = P_{k/k-1} \tag{85}$$

$$M^{-1}_{k/k} = M^{-1}_{k/k-1} + \frac{1}{\sigma_\varepsilon^2} z_k' z_k \tag{86}$$

where

$$z_k = \left[y_k , \bar{u}_k \right] \tag{87}$$

The above problem is solved via the OLFO method and the following number is computed.

$$V(y^k, \bar{u}_k) = \bar{y}^{\,2}_{k+1} + r\bar{u}^2_k + V_{OLFO}(y^k, \bar{y}_{k+1}) \tag{88}$$

Now a new value for $\bar{u}_k$ is chosen and the whole procedure is repeated. The usual procedure is to start with the CE control and then search in the neighborhood so as to find a better control. The control minimizing $V(y^k, \bar{u}_k)$ is applied and the method is started anew in the next time step.

The method has at least one advantage, namely that it guarantees a better control than the starting one which can be the CE one. Tse and Bar-Shalom (1973) have shown numerical results in which this method was better than CE by one order of magnitude.

The main disadvantage of it is that in general it involves a search in a m-dimensional space, where m is the dimension of the control vector. Unless the control space is bounded, this search will result in a local minimum of $V(y^k)$ with respect to u_k. In addition, as was seen in section 6.4, the exact OLFO control is hard to find and approximations might be used. In such cases the quantity V_{OLFO} in (88) is substituted by an approximate one. Therefore, the minimization of (88) with respect to $\bar{u}_k$ will be an approximate one.

Modifications of this method are easy to visualize. One which seems to us particularly appealing is to substitute for V_{OLFO} in (88) the quantity V_{S1}, namely the cost computed with the S1 method analyzed in section 6.3. Without some numerical studies it is quite difficult to assert a-priori which method would perform better.

The dual nature of the one-measurement-optimal feedback method is manifested by the fact that the covariances of the parameters at time k+1 are explicit functions of the control applied at time k. The dependence of the future covariances on the present control is nonlinear and quite complicated. Thus since it is hard to compute the explicit dependence analytically numerical evaluations have to be made. For on line applications this can be quite costly.

7.2 Adaptive Covariance Method.

This quite interesting method was proposed by McRae (1972). Here we shall present the main idea, and we shall extend her results to our problem, and give them a shape suitable for numerical computation, which she has not done.

Suppose we are at time k and we have observed y^k and u^{k-1}. We would like to choose $u_k, u_{k+1}, \ldots, u_{N-1}$ so as to minimize the quantity

$$V(y^k) = E\{\sum_{i=k}^{N-1} (y_{i+1}^2 + ru_i^2)/y^k, u^{k-1}\} \equiv$$
$$E\{\sum_{i=k}^{N-1} (y_{i+1}^2 + ru_i^2)/k-1\} \tag{89}$$

subject to

$$y_{j+1} = z_j p_j + g_j \qquad j \geq k \tag{90}$$

$$p(p_k/k-1) = N(p_{k/k-1}, M_{k/k-1}) \tag{91}$$

Our future knowledge of the parameters p_j will be governed by the posterior density of p_j given future data. From section 3 we know that the future posterior densities of p_j will be normal with means and covariances evolving by the formulas:

$$p_{j/j} = M_{j/j}\left[M^{-1}_{j/j-1}p_{j/j-1} + \frac{1}{\sigma^2_\varepsilon} z'_j y_{j+1}\right] \tag{92}$$

$$M^{-1}_{j/j} = M^{-1}_{j/j-1} + \frac{1}{\sigma^2_\varepsilon} z'_j z_j \tag{93}$$

$$p_{j/j-1} = p_{j-1/j-1} \tag{94}$$

$$M_{j/j-1} = M_{j-1/j-1} + R \tag{95}$$

for $j \geq k$ with initial conditions given in (91).

In view of (90) the constraints (92)-(93) are stochastic. We make the following approximation similar to McRae's, that renders them deterministic. We assume that the evolution of means and covariances will be deterministic and given by the following formulas.

$$p_{j/j} \simeq M_{j/j}\left[M^{-1}_{j/j-1}p_{j/j-1} + E(\frac{1}{\sigma^2_\varepsilon} z'_j y_{j+1} /k-1)\right] \tag{96}$$

$$M^{-1}_{j/j} \simeq M^{-1}_{j/j-1} + E\left[\frac{1}{\sigma^2_\varepsilon} z'_j z_j /k-1\right] \tag{97}$$

$$p_{j/j-1} = p_{j-1/j-1} \tag{98}$$

$$M_{j/j-1} = M_{j-1/j-1} + R \tag{99}$$

Thus the future means and covariances are functions of quantities that are to be calculated at time k, i.e. $u_k, u_{k+1}, \ldots, u_{N-1}$.

Let us analyze (96) a little further.

$$E\left(\frac{1}{\sigma_\varepsilon^2} z'_j y_{j+1} / k-1\right) = \frac{1}{\sigma_\varepsilon^2} E\left[z'_j E(y_{j+1} / y^j, u^j) / k-1\right] =$$

$$\frac{1}{\sigma_\varepsilon^2} E\left[z'_j z_j E(p_j / y^j, u^j) / k-1\right] =$$

$$\frac{1}{\sigma_\varepsilon^2} E\left[z'_j z_j p_{j/j-1} / k-1\right] \qquad (100)$$

Since $p_{j/j-1}$ is deterministic it can be factored out of (100). Therefore, (96) becomes

$$p_{j/j} = M_{j/j}\left[M^{-1}_{j/j-1} + E\left(\frac{1}{\sigma_\varepsilon^2} z'_j z_j / k-1\right)\right] p_{j/j-1} \qquad (101)$$

Equation (97) now implies that

$$p_{j/j} = p_{j/j-1} = p_{j-1/j-1} = \ldots = p_{k/k-1} \qquad (102)$$

Thus implicit in assumptions (96)-(99) is the fact that the future mean is not affected by the controls but the future covariance is via (97). The problem that we solve is the following.

Minimize $V(y^k)$ in (89) with respect to $u_k, u_{k+1}, \ldots, u_{N-1}$ subject to the stochastic constraints (90), and given that the future densities of the parameters have means given by (102) and covariances by (97). The problem that we pose is both stochastic and deterministic because half the constraints are stochastic, namely (90), and half deterministic, namely (97). We solve it, following McRae, by applying dynamic programming to a criterion which is (89) augmented by products of the deterministic constraints and deterministic Lagrange multipliers.

The complete analysis is given in appendix A. The result is that the controls $u_k, u_{k+1}, \ldots, u_{N-1}$ are linear functions of $y_k, y_{k+1}, \ldots, y_{N-1}$ respectively with gains given by the solution of a two-point-boundry-value (TPBV) problem. The complete set of equations is the following (For proof see appendix A).

$$u_j = -G_j^{-1} F_j y_j \tag{103}$$

$$G_j = r + (1 + K_{j+1})(b^2_{j/j-1} + M^b_{j/j-1}) - \frac{1}{\sigma_\varepsilon^2} L^b_j \tag{104}$$

$$F_j = (1 + K_{j+1})(a_{j/j-1} b_{j/j-1} + M^{ab}_{j/j-1}) - \frac{1}{\sigma_\varepsilon^2} L^{ab}_j \tag{105}$$

$$K_j = (1 + K_{j+1})(a^2_{j/j-1} + M^a_{j/j-1}) - \frac{1}{\sigma_\varepsilon^2} L^a_j - G_j^{-1} F_j^2 \tag{106}$$

$$L_j = (I + RM^{-1}_{j/j})^{-1} L_{j+1} (I + M^{-1}_{j/j} R)^{-1} - M_{j/j} P_{j+1} M_{j/j} x_{j+1} \tag{107}$$

$$P_j \equiv \begin{bmatrix} 1 & -G_j^{-1} F_j \\ -G_j^{-1} F_j & G_j^{-2} F^2 \end{bmatrix} \tag{108}$$

$$x_j \equiv E(y_j^2/k-1) = \sigma_\varepsilon^2 + x_{j-1}\, tr\, \{P_{j-1}(p_{j-1/j-2}\, p'_{j-1/j-2} + M_{j-1/j-2})\} \quad (109)$$

$$P_{j/j} \equiv \begin{bmatrix} a_{j/j} \\ b_{j/j} \end{bmatrix} = P_{j/j-1} = P_{j-1/j-1} = \cdots = P_{k/k-1} \quad (110)$$

$$M^{-1}_{j/j} = M^{-1}_{j/j-1} + \frac{1}{\sigma_\varepsilon^2} P_j x_j \quad (111)$$

$$M_{j/j-1} = M_{j-1/j-1} + R \quad (112)$$

The boundary conditions are $K_N = 0$, $L_{N-1} = 0$, $x_k = y_k^2$, $P_{k/k-1}$ and $M_{k/k-1}$ known.

The solutions of the above equations must be carried at each step k and only u_k applied to the system. Then a new measurement is taken and the procedure must be repeated. What is interesting about this method is that the future controls are linear and influence all the future covariances. We have not as yet examined numerical ways to solve the above TPBV problem.

7.3 Two-Step Optimal Adaptive Control.

This method, to our knowledge has not been suggested before. The idea is the following. Assume that we are at time k having observed y^k, u^{k-1}. Then assume that optimization is to be done only for two more periods. Also assume that the one future value of the parameter b is and equal to $b_{k/k-1}$. Then carry out the two-step backward dynamic programming recursion. The assumption that b_{k+1} is constant and equal to $b_{k/k-1}$ is sufficient to render the minimization with respect to u_k equivalent to minimization of a quadratic function of u_k.

$$V^*(y^k) = \min_{u_k} E\left[y_{k+1}^2 + ru_k^2 + V^*(y^{k+1})/k-1\right] \tag{113}$$

where

$$V^*(y^{k+1}) = \min_{u_{k+1}} E\left[y_{k+2}^2 + ru_{k+1}^2/k\right] \tag{114}$$

At time k+1 the minimization (114) is straightforward. We obtain

$$u_{k+1}^* = -G_{k+1}y_{k+1} \tag{115}$$

where

$$G_{k+1} = (r + b_{k+1}^2)^{-1}b_{k+1}a_{k+1/k} \tag{116}$$

$$V^*(y^{k+1}) = Hy_{k+1}^2 + F \tag{117}$$

where

$$H = a^2_{k+1/k} + M^a_{k+1/k} - (r + b^2_{k+1})^{-1} b^2_{k+1} a^2_{k+1/k} \quad (118)$$

$$F = \sigma^2_\varepsilon \quad (119)$$

At time k we have

$$V^*(y^k) = \min_{u_k} E\left[(1 + H)y^2_{k+1} + ru^2_k + F/k-1\right] \quad (120)$$

From section 3 we know that

$$a_{k+1/k} = a_{k/k} = a_{k/k-1} + M^a_{k/k-1} y_k (y^2_k M^a_{k/k-1} + \sigma^2_\varepsilon)^{-1} (y_{k+1} - a_{k/k-1} y_k - b_{k/k-1} u_k) \quad (121)$$

$$M^a_{k+1/k} = M^a_{k/k} + \sigma^2_\theta = M^a_{k/k-1} - M^{a^2}_{k/k-1} y^2_k (y^2_k M^a_{k/k-1} + \sigma^2_\varepsilon)^{-1} + \sigma^2_\theta \quad (122)$$

If we substitute for y_{k+1} in (121) the innovation becomes

$$(a_k - a_{k/k-1})y_k + (b_k - b_{k/k-1})u_k + \varepsilon_k \quad (123)$$

We make the assumption

$$(b_k - b_{k/k-1}) = 0 \quad (124)$$

which is what will render the problem tractable.

Of course, if b_k is a-priori known then the assumption (124) will be a true fact, and not an approximation.

By substituting for $a_{k+1/k}, M^a_{k+1/k}$ in H, via (121)-(122) and substituting for y_{k+1} in (120) we arrive at an expression whose conditional expectation is easy to take. In addition the resulting expression is quadratic in u_k. The calculations are lengthy but straightforward and they are shown in appendix C. The optimizing u_k is

$$u^*_k = -D^{-1}_k f_k \tag{125}$$

where

$$D_k = r + (1 + M^a_{k+1/k})(b^2_{k/k-1} + M^b_{k/k-1}) + \frac{r}{r + b^2_{k/k-1}} \{a^2_{k/k-1} \cdot$$

$$\cdot (b^2_{k/k-1} + M^b_{k/k-1}) + X^2_k \Big[E\left[b^2_k(a_k - a_{k/k-1})^2/k-1\right] y^2_k +$$

$$\sigma^2_\varepsilon (b^2_{k/k-1} + M^b_{k/k-1}) \Big] \} \tag{126}$$

$$f_k = y_k \{ (1 + M^a_{k+1/k})E(a_k b_k/k-1) + \frac{r}{r + b^2_{k/k-1}} \Big[a^2_{k/k-1}E(a_k b_k/k-1) +$$

$$X^2_k \left[y^2_k E\left[a_k b_k(a_k - a_{k/k-1})^2/k-1\right] + \sigma^2_\varepsilon E(a_k b_k/k-1)\right] +$$

$$2a_{k/k-1}X_k y_k E\left[a_k b_k(a_k - a_{k/k-1})k-1\right]\} +$$

$$\frac{r}{r + b^2_{k/k-1}} \left[2X^2_k y_k E\left[b_k(a_k - a_{k/k-1})/k-1\right]\sigma^2_\varepsilon + 2a_{k/k-1}b_{k/k-1}X_k\sigma^2_\varepsilon\right] \quad (126)$$

where

$$X_k = M^a_{k/k-1}y_k(y^2_k M^a_{k/k-1} + \sigma^2_\varepsilon)^{-1} \quad (127)$$

The control u_k is thus a highly nonlinear function of y_k. We can also see that even if we make the assumption that b_k is equal to $b_{k-1/k-2}$ it is impossible to carry out one more recursive dynamic programming step because of the complex nonlinear dependence of $V^*(y^k)$ on y_k.

This control law is dual and it takes into account future adaptation of the mean but not the variance of a. It is quite simple to compute since it does not involve the solution of any iterative system of equations like the previous methods.

8. NUMERICAL COMPARISONS

In this section we show the results of some initial Monte Carlo comparisons of all the non-dual methods mentioned before except the OLFO one, for which exact computations are tedious as seen in section 6.4. and inexact computations give strange results (cf. Ku and Athans (1973)).

The methods compared are denoted by the following initials:

T - Control with perfectly known parameters (cf. section 5)

CE - Certainty equivalence method (cf. section 5)

W - Wouters' method (cf. section 6.1)

WW - Wieslander's and Wittenmark's method (cf. section 6.2)

WM - Modified WW (cf. equation (49))

S1 - Sequential stochastic control (cf. section 6.3)

For all the methods except T, which does not involve learning, the parameter updating was done with the Kalman filter analyzed in section 3.

We now state the results for four experiments that were conducted. Table 1 summarizes the conditions of each experiment. The first column denotes a code name for the experiment. The second column denotes a code name for the true parameters used in generating the data. The third column lists the covariances of the system error. The random numbers that were created had the indicated covariances and were normal. The M_o column lists the initial covariance matrix of the parameters. For every run the initial values of the parameters were chosen by random sampling from a normal density with mean $\bar{p}_o$, listed in the last column, and covariance matrix M_o. The column labeled R lists the covariance matrices used for the error terms in the parameter equations (cf. (7) and

Table 1. List of Monte Carlo Experiments

EXPERIMENT	TRUE COEF.	σ_ε^2	M_o	R	No. of RUNS	Y_o	r	$\bar{p}_o'$
E_1	AB1	.25	$10^{-2}I_2$	$10^{-3}I$	20	3	.3	(.8,.2)
E_2	AB2	.25	$10^{-2}I_2$	$\begin{bmatrix} .09 & 0 \\ 0 & .01 \end{bmatrix}$	20	3	.3	(-.63,.083)
E_3	AB3	.25	$10^{-2}I_2$	$10^{-2}I$	20	3	.3	(.8,.3)
E_5	AB4	.25	$10^{-2}I_2$	$10^{-2}I$	20	3	.3	(.6,-.2)

Table 2. List of Average Cost in Experiments

METHOD: EXPERIMENT	T	CE	W	WW	WM	S1
E_1	20.158	20.237	93.444	61.258	20.507	20.189
E_2	29.524	625.557	5185.37	23584.1	3.386×10^9	64009.4
E_3	19.661	20.921	272.849	31.333	21.997	20.76
E_5	16.412	17.99	47.485	27.421	20.552	18.257

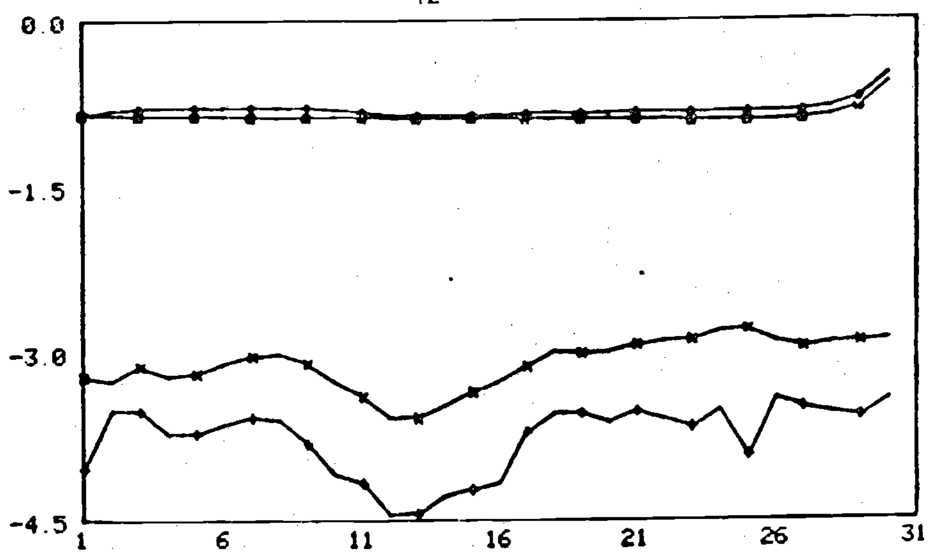
(10)). The remaining three columns list the number of runs, the initial value of y_o and the control penalty r respectively. All runs were for 30 periods.

In experiment E1 the true parameter a_t was constant and the true b_t was a slow trend. In E2 the true a_t and b_t were generated using equation (7) with initial values (-.63,.083), as shown in the last column of Table 1, and normal random errors with zero means and covariances $\sigma_\theta^2 = .09$ and $\sigma_\eta^2 = .01$. In E3 the true parameters were both time varying with some trends and sudden jumps. In E5 both the parameters were constant with a_t equal to .7 and b_t equal to -.4.

In Table 2 we show the average cost for the 20 runs. The first thing that we notice is that the CE method performs quite well, surpassed at some experiments only by S1. We see that the W and WW methods which are minimum variance ones involve excessively high control cost. In experiment E2 the parameter a_t was unstable for half of the controlling period, and we see that all suboptimal methods perform poorly. This is a disturbing fact and was also observed by Ku and Athans (1973) in their simulations of the OLFO method.

Figures 1-12 show the average control gains and the average parameter estimated resulting from the 20 Monte Carlo runs of each experiment. It is interesting to notice that for E2 in which, as seen in Table 2, none of the methods gave good controls, nevertheless the estimates of the parameters are quite satisfactory. In general W, WW, and WM give the worst results with CE and S1 always superior to those three. The experiments, however, did not result in a distinct ordering of CE and S1.

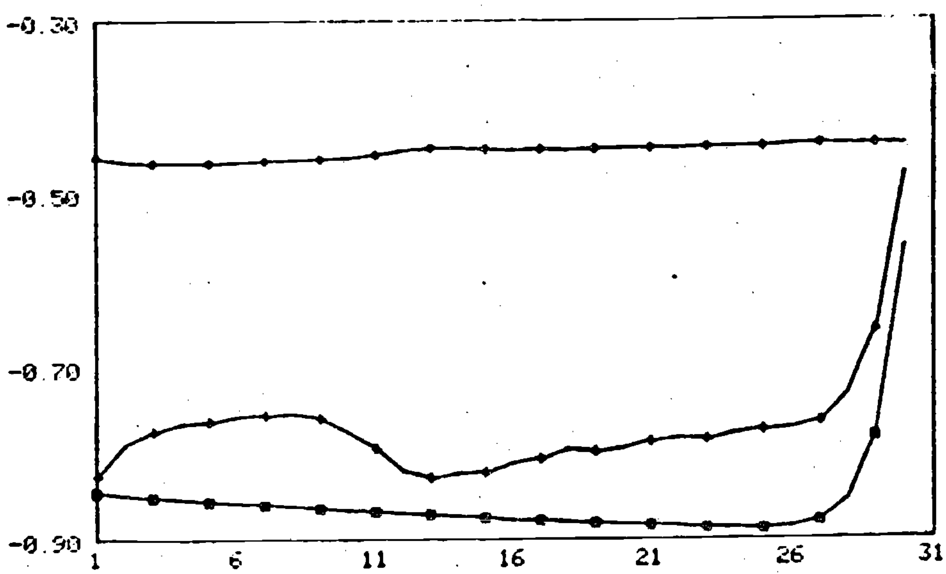
There is still a lot of work to be done in comparing these methods and comparing them with the dual methods described in section 8. The dual methods should give better results than the non-dual ones. On the other hand the dual ones are all, with the exception of the one described in section 7.3, quite costly.



TIME BOUNDS: 1 TO 30

SYMBOL SCALE NAME

■	#1	E1_TG
●	#1	E1_CG
◆	#1	E1_NG
✖	#1	E1_NNG

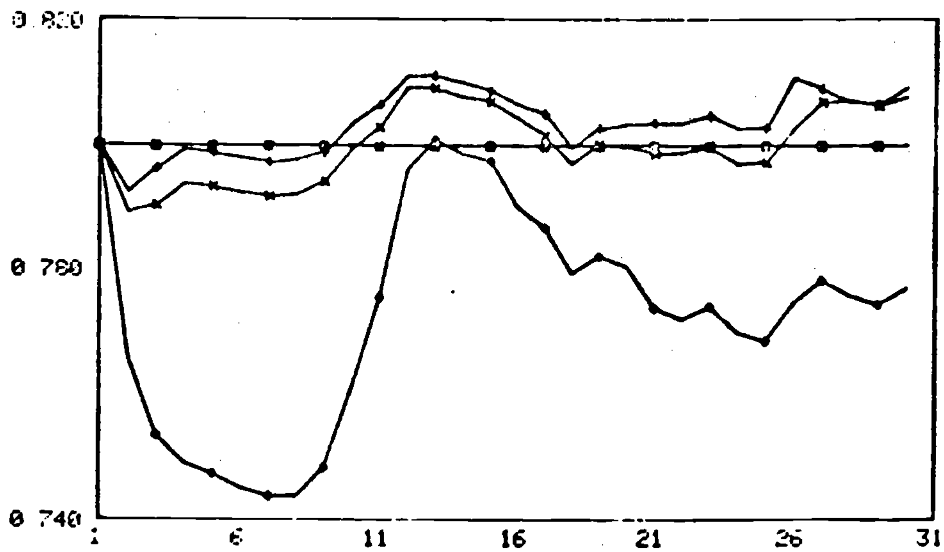


TIME BOUNDS 1 TO 30

SYMBOL SCALE NAME

■	#1	E1_TG
●	#1	E1_NNG
◆	#1	E1_SIG

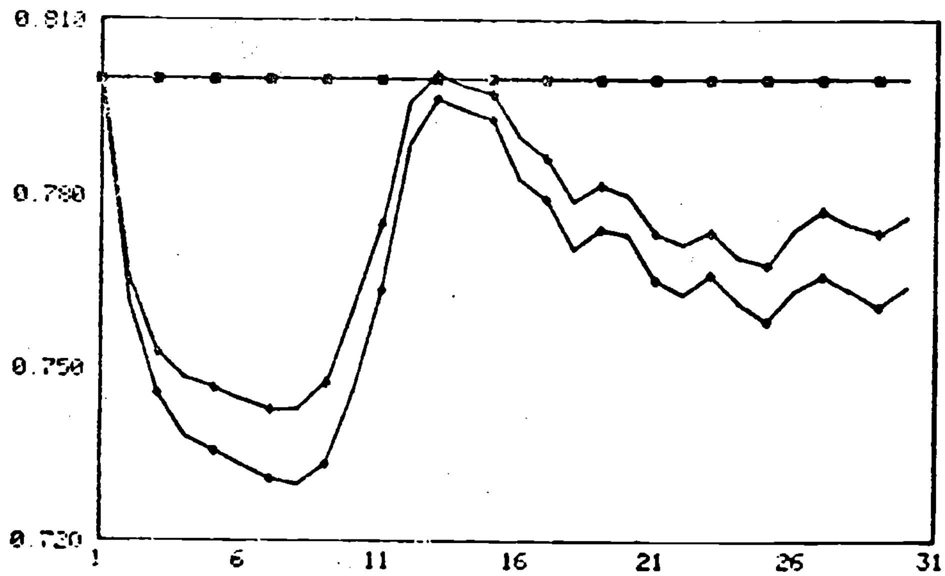
Figure 1. Control Gains for E1.



TIME BOUNDS 1 TO 30

SYMBOL SCALE NAME

■	#1	AB1_C1
●	#1	E1_CPC_C1
◆	#1	E1_WPC_C1
✱	#1	E1_WWPC_C1

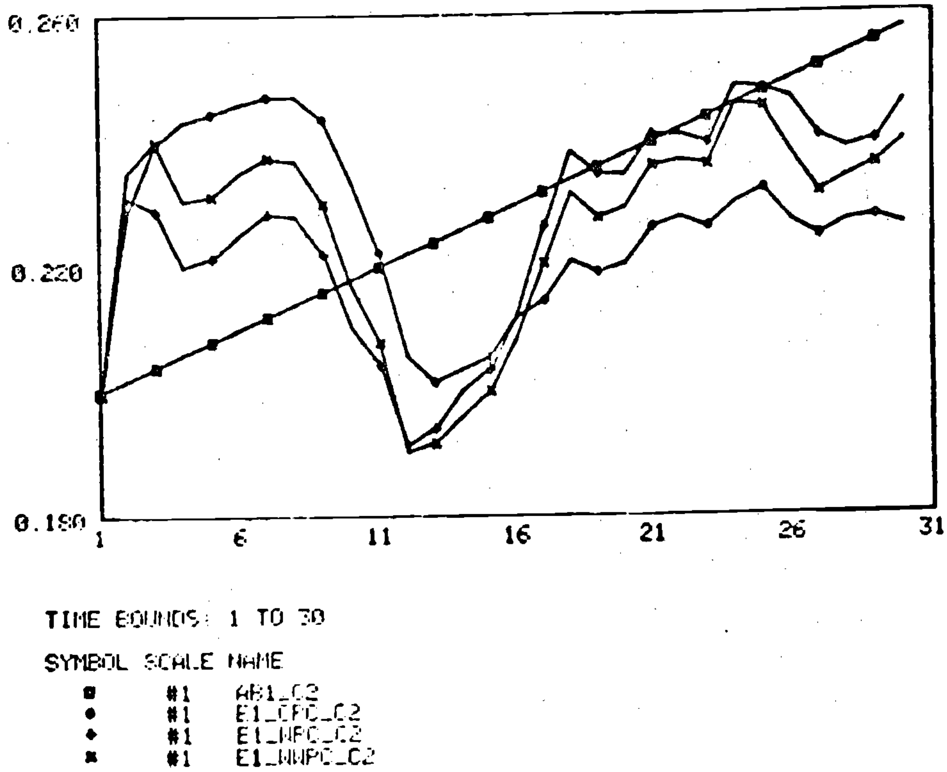


TIME BOUNDS 1 TO 30

SYMBOL SCALE NAME

■	#1	AB1_C1
●	#1	E1_WMPC_C1
◆	#1	E1_SIPC_C1

Figure 2. Estimates of a_t in E1.



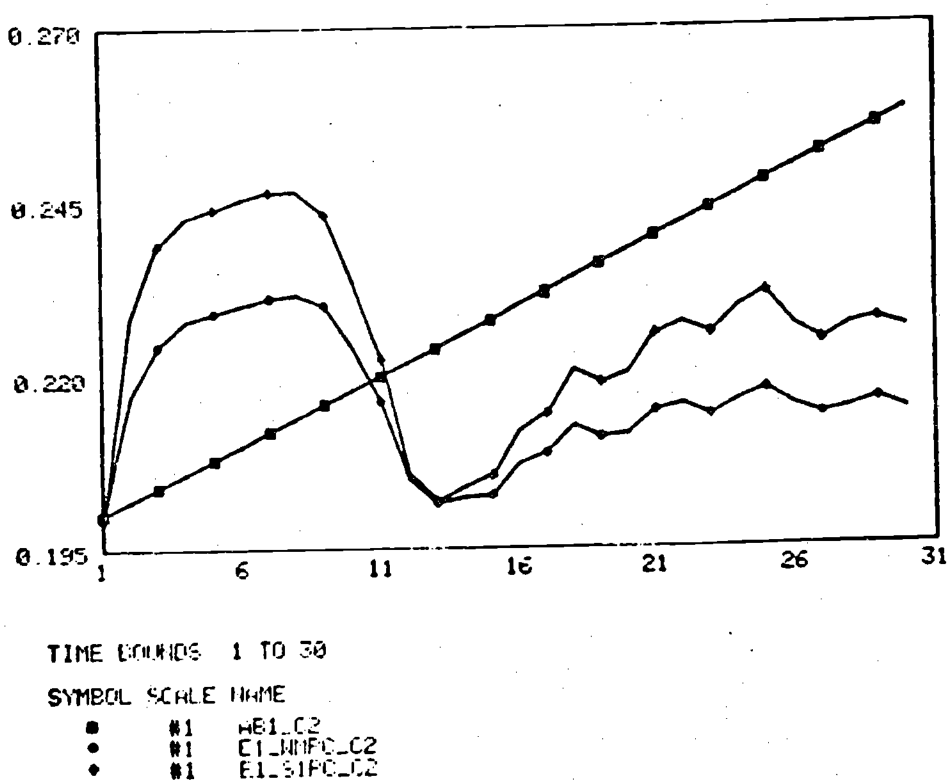
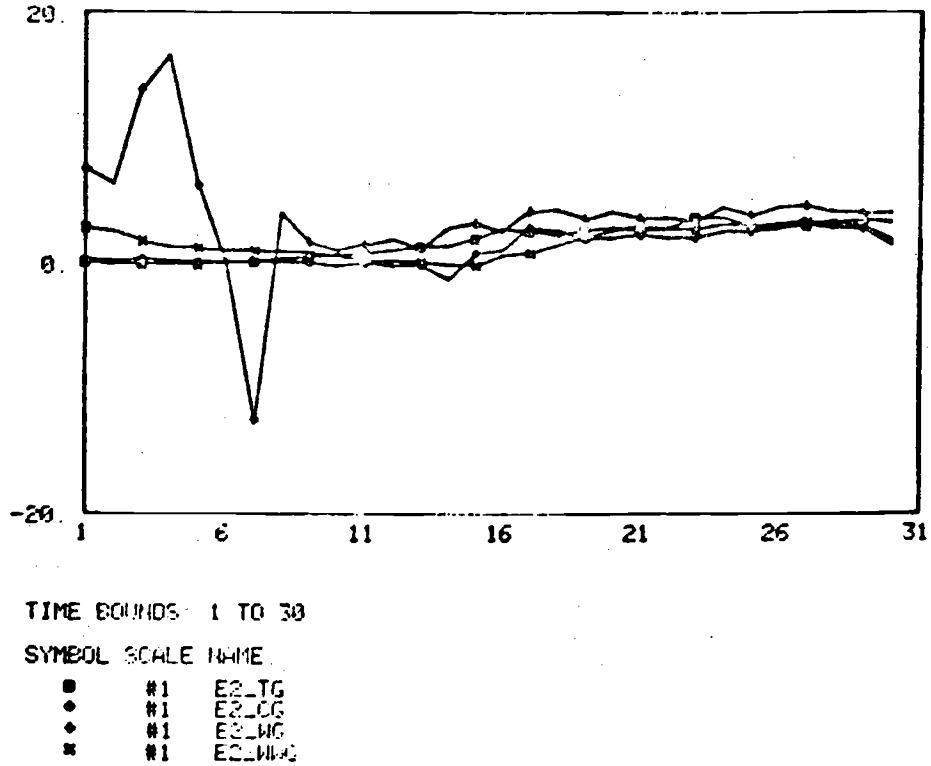


Figure 3. Estimates of b_t in E1.



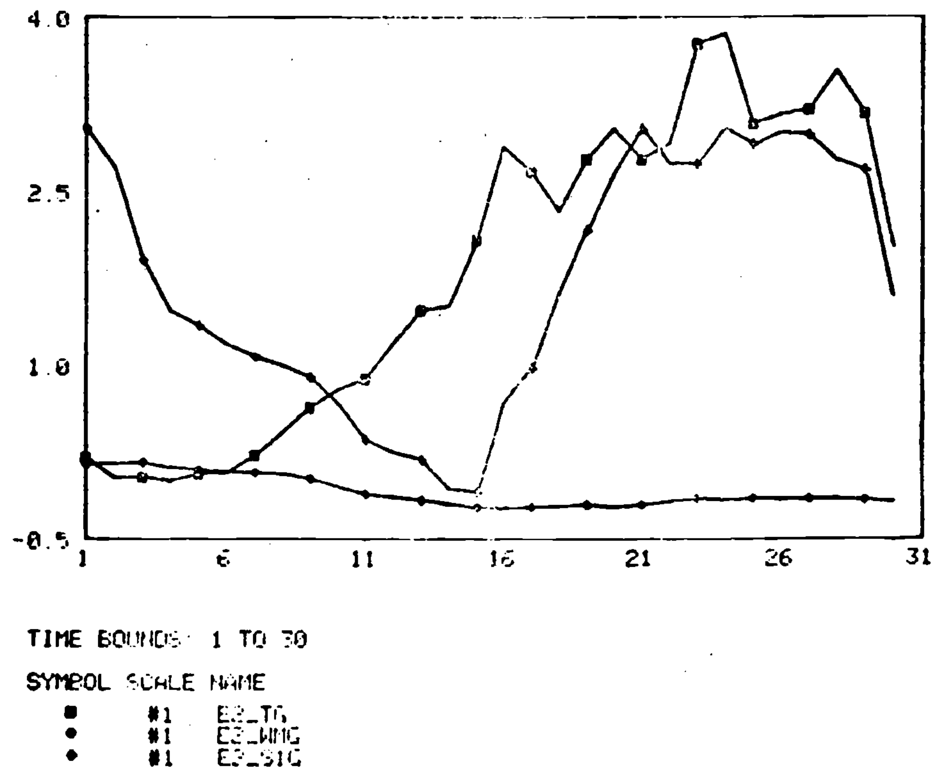
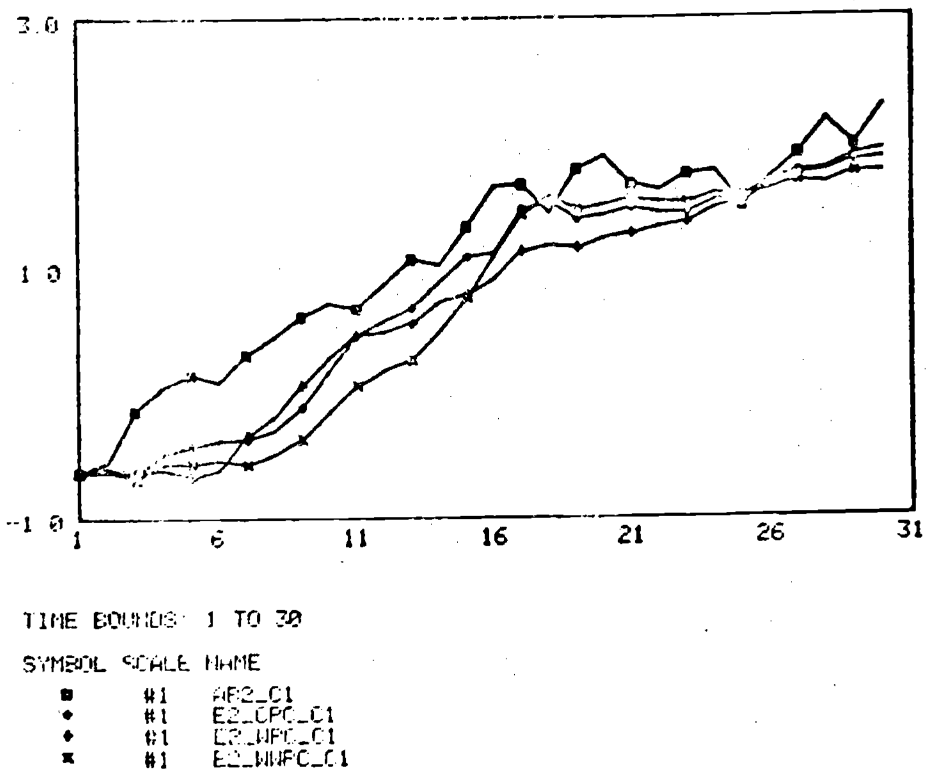


Figure 4. Control Gains for E2.



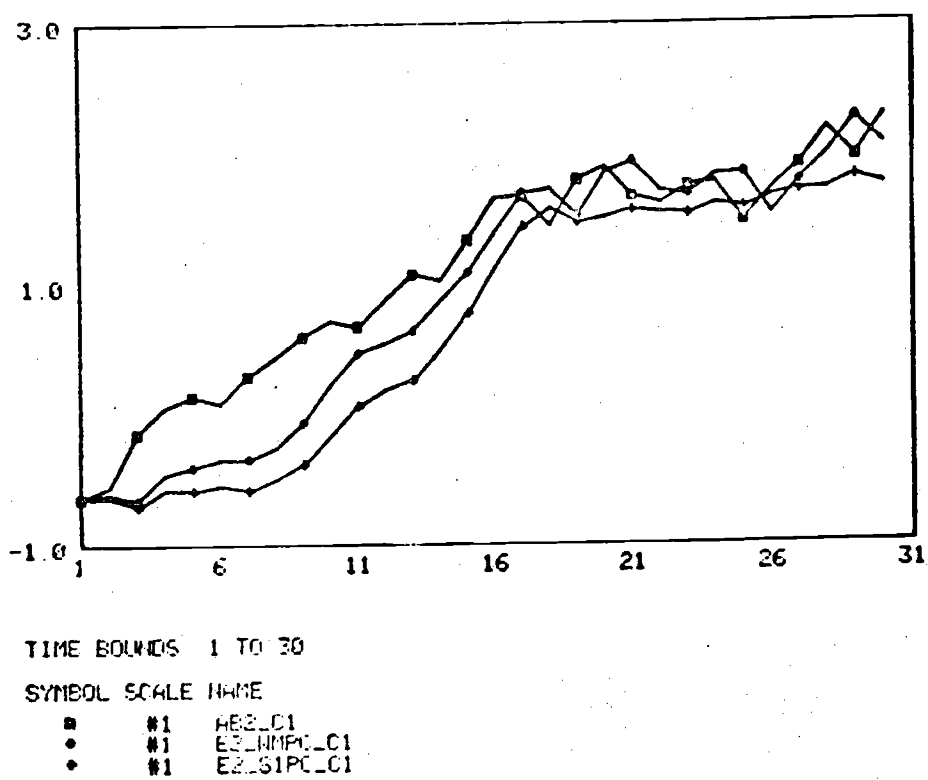
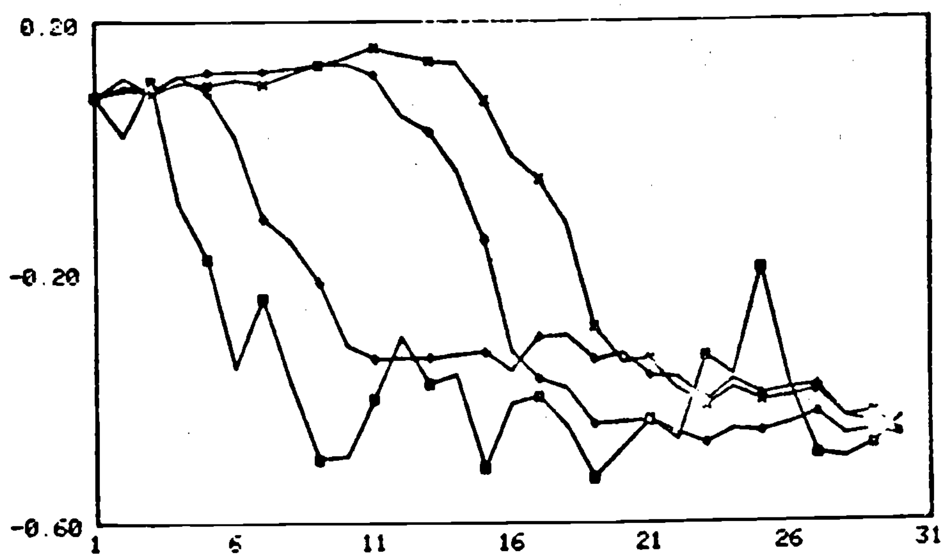


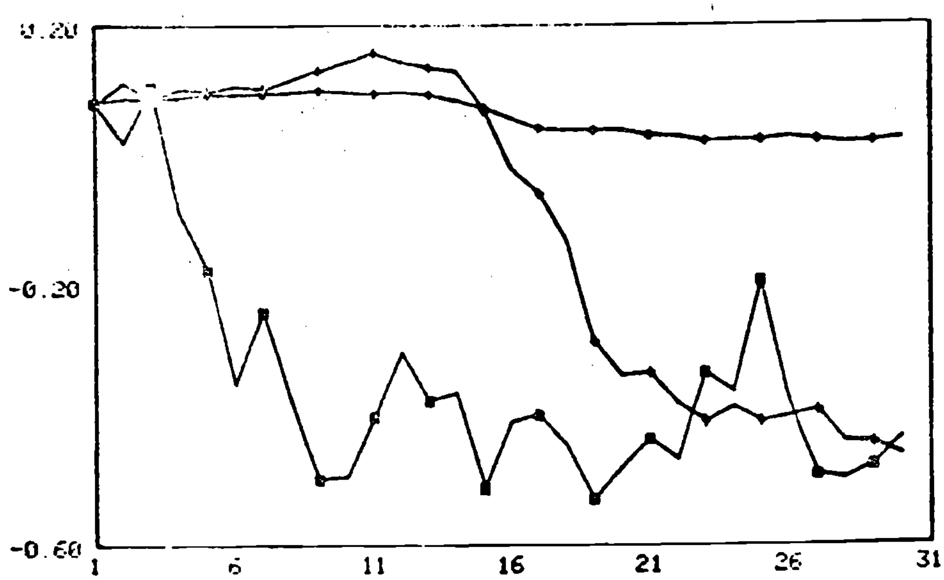
Figure 5. Estimates of a_t in E2.



TIME BOUNDS 1 TO 30

SYMBOL SCALE NAME

- ■ #1 AB2_C2
- ● #1 E2_CPC_C2
- ◆ #1 E2_WPC_C2
- ✱ #1 E2_WMPC_C2

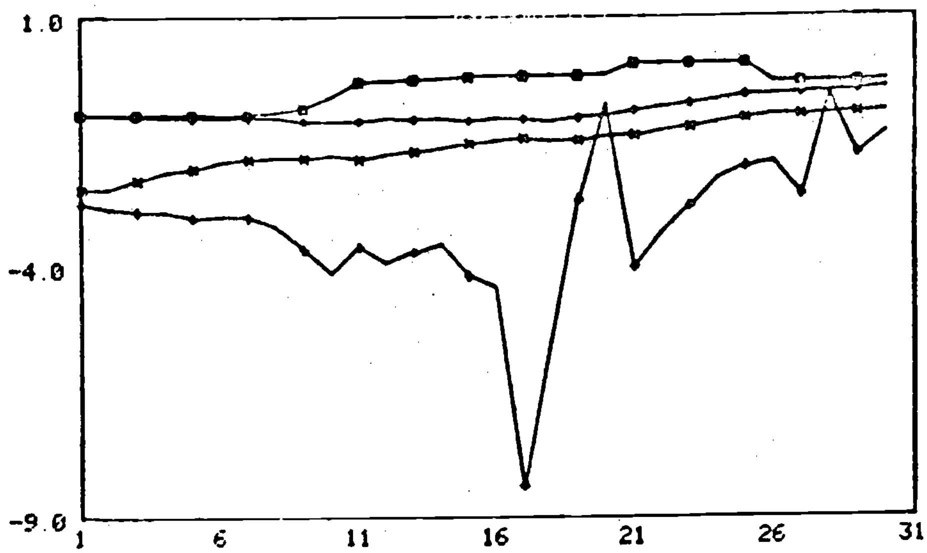


TIME BOUNDS 1 TO 30

SYMBOL SCALE NAME

- ■ #1 AB2_C2
- ● #1 E2_WMPC_C2
- ◆ #1 E2_S1PC_C2

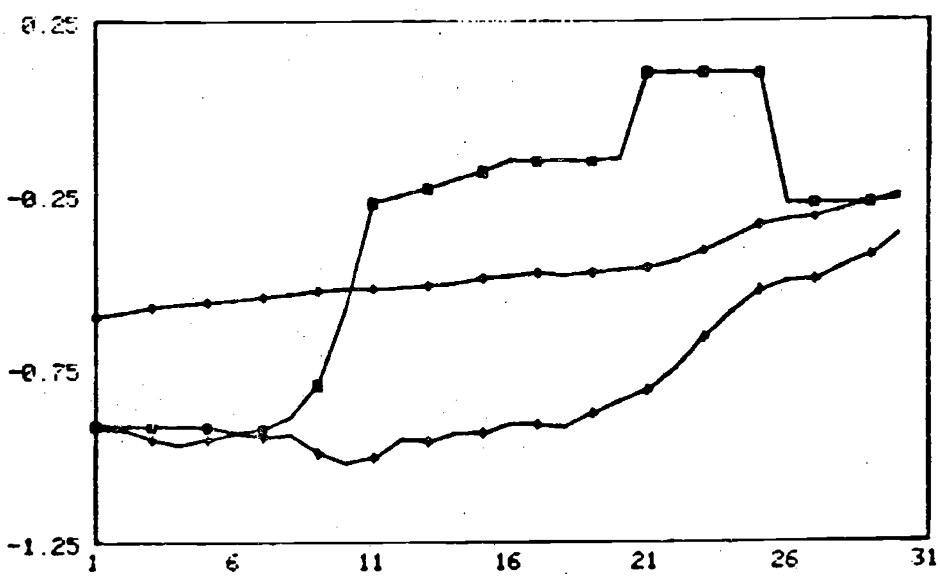
Figure 6. Estimates of b_t in E2.



TIME BOUNDS 1 TO 30

SYMBOL SCALE NAME

SYMBOL	SCALE	NAME
■	#1	E3_TG
•	#1	E3_CG
♦	#1	E3_WG
✱	#1	E3_WWG



TIME BOUNDS 1 TO 30

SYMBOL SCALE NAME

SYMBOL	SCALE	NAME
■	#1	E3_TG
•	#1	E3_WWG
♦	#1	E3_S1G

Figure 7. Control Gains for E3.

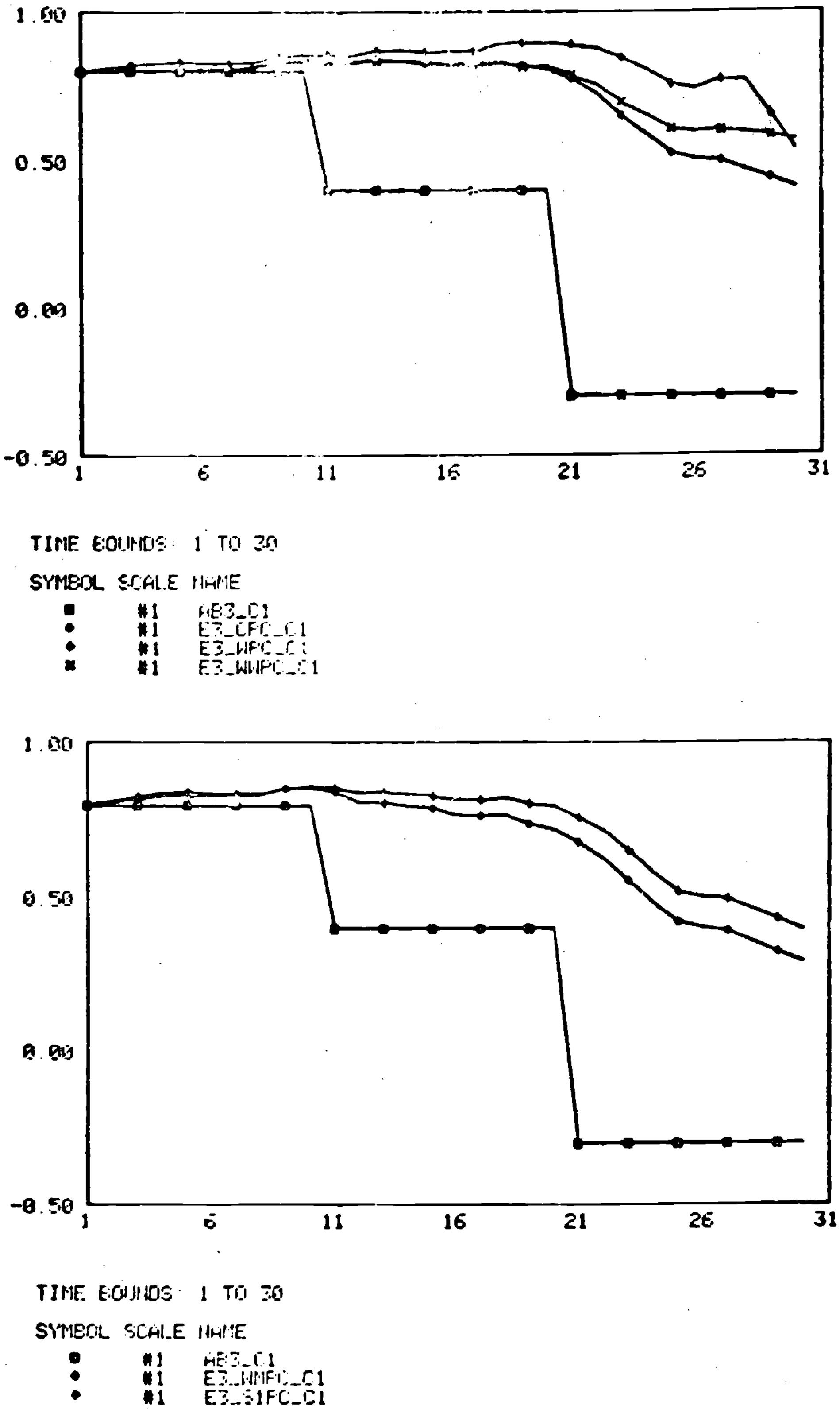
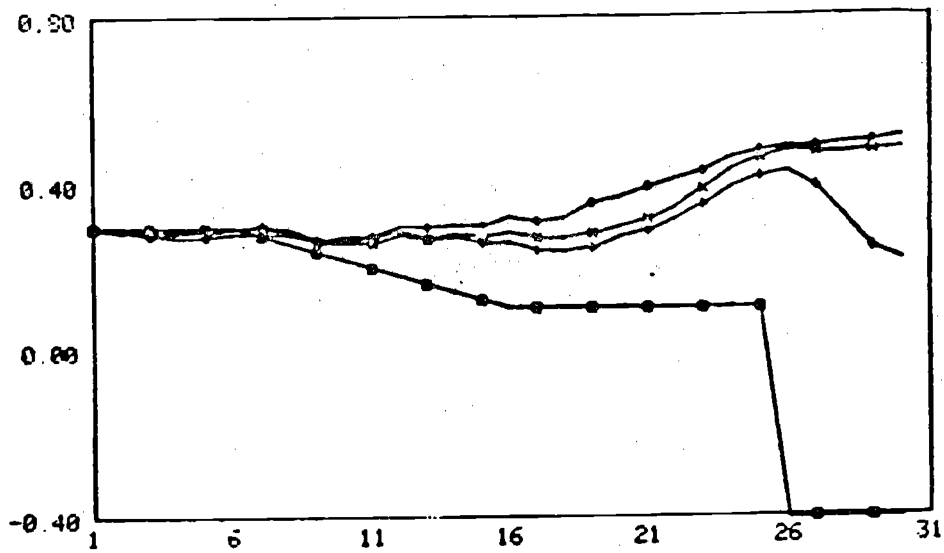


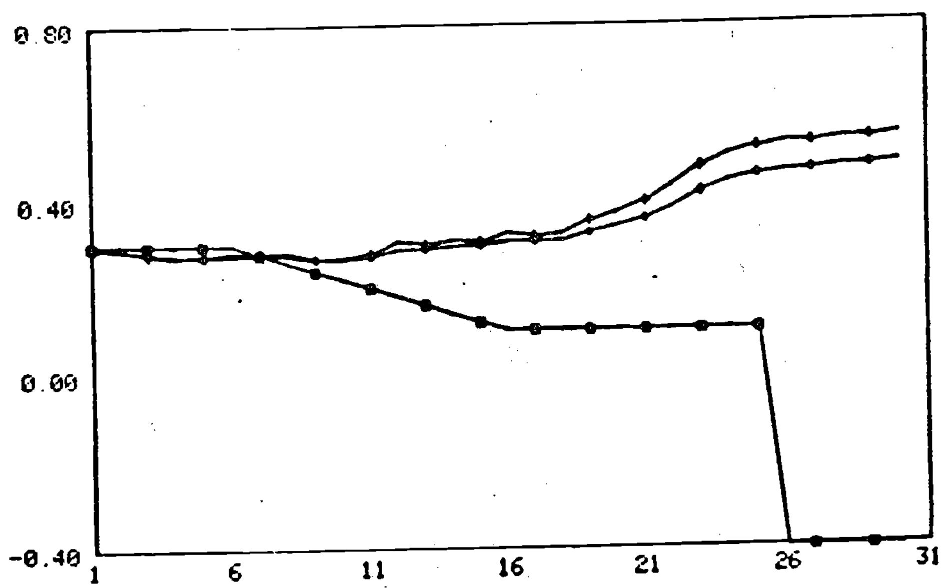
Figure 8. Estimates of a_t in E3.



TIME BOUNDS: 1 TO 30

SYMBOL SCALE NAME

SYMBOL	SCALE	NAME
■	#1	AB3_C2
•	#1	E3_CPC_C2
◆	#1	E3_WPC_C2
✱	#1	E3_WWPC_C2



TIME BOUNDS: 1 TO 30

SYMBOL SCALE NAME

SYMBOL	SCALE	NAME
■	#1	AB3_C2
•	#1	E3_WMPC_C2
◆	#1	E3_S1PC_C2

Figure 9. Estimates of b_t in E3.

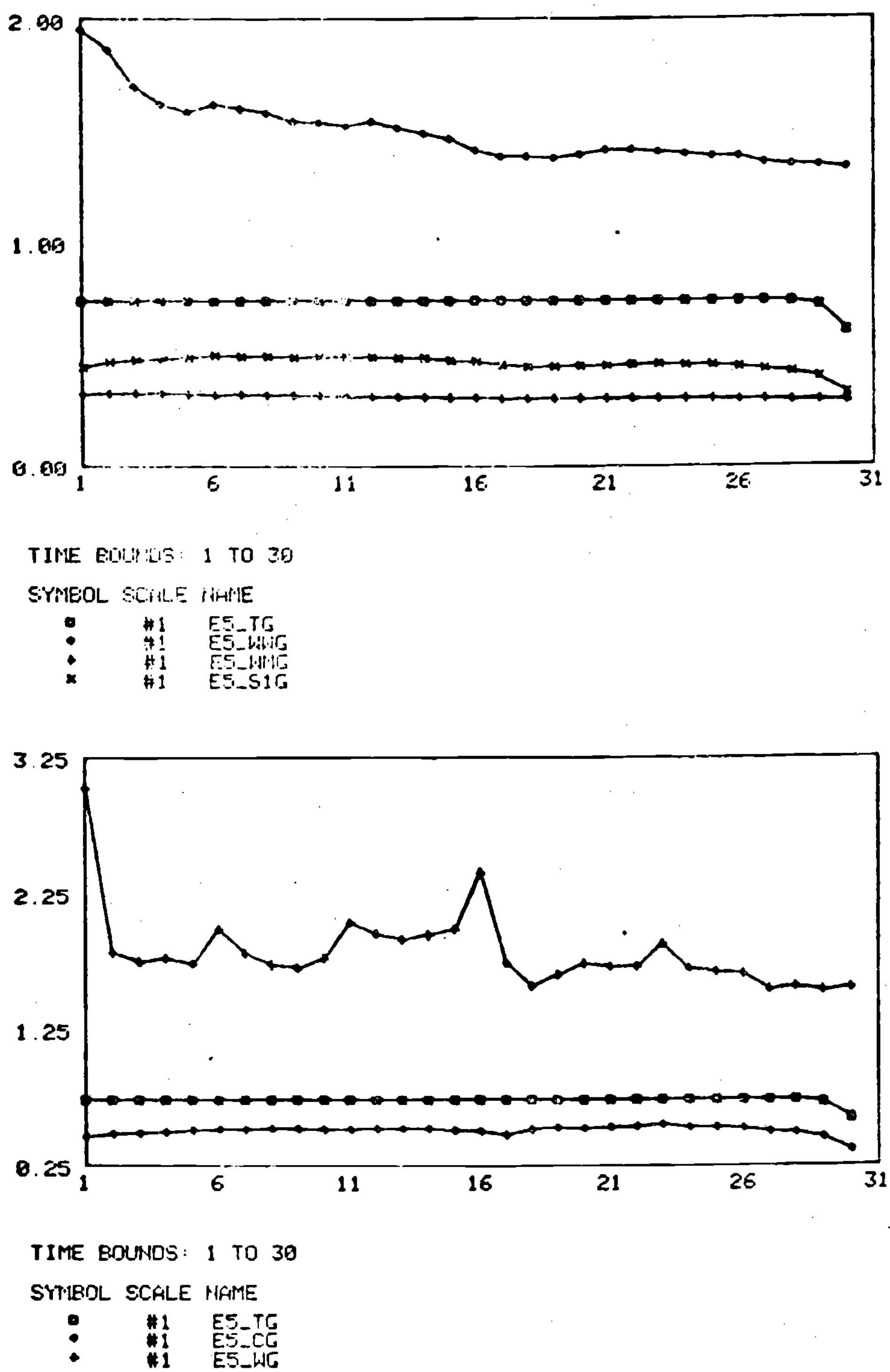
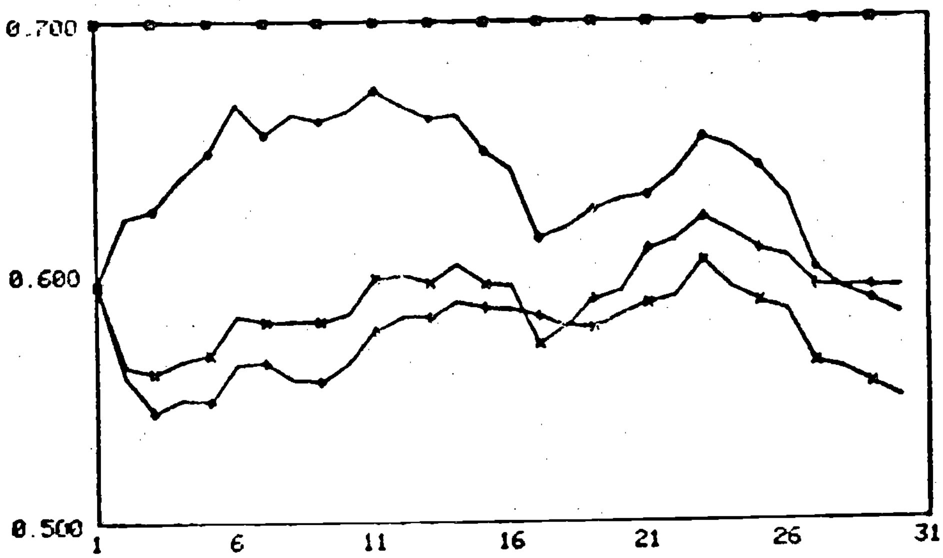


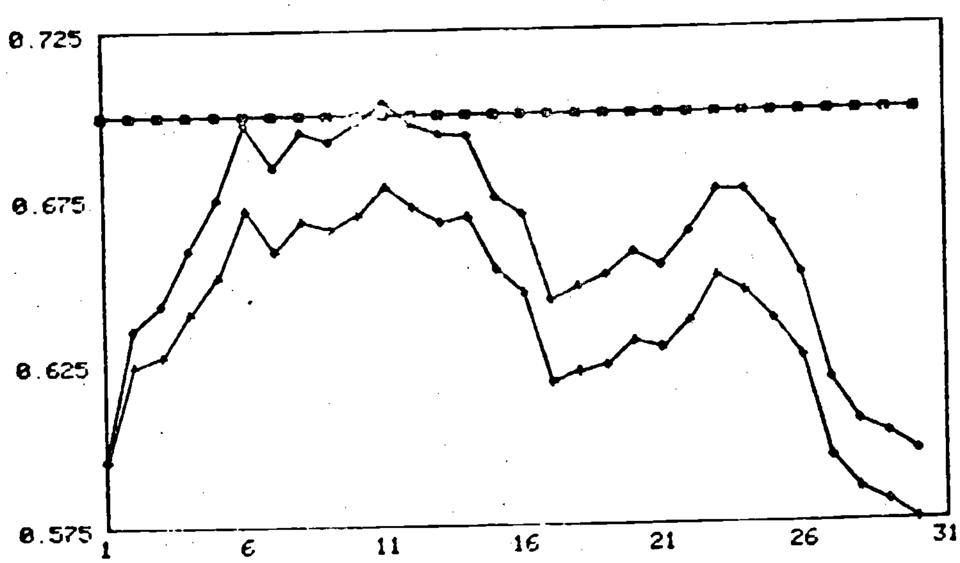
Figure 10. Control Gains for E5.



TIME BOUNDS: 1 TO 30

SYMBOL SCALE NAME

■ #1 AB4_C1
• #1 E5_CPC_C1
♦ #1 E5_WPC_C1
✱ #1 E5_WWPC_C1

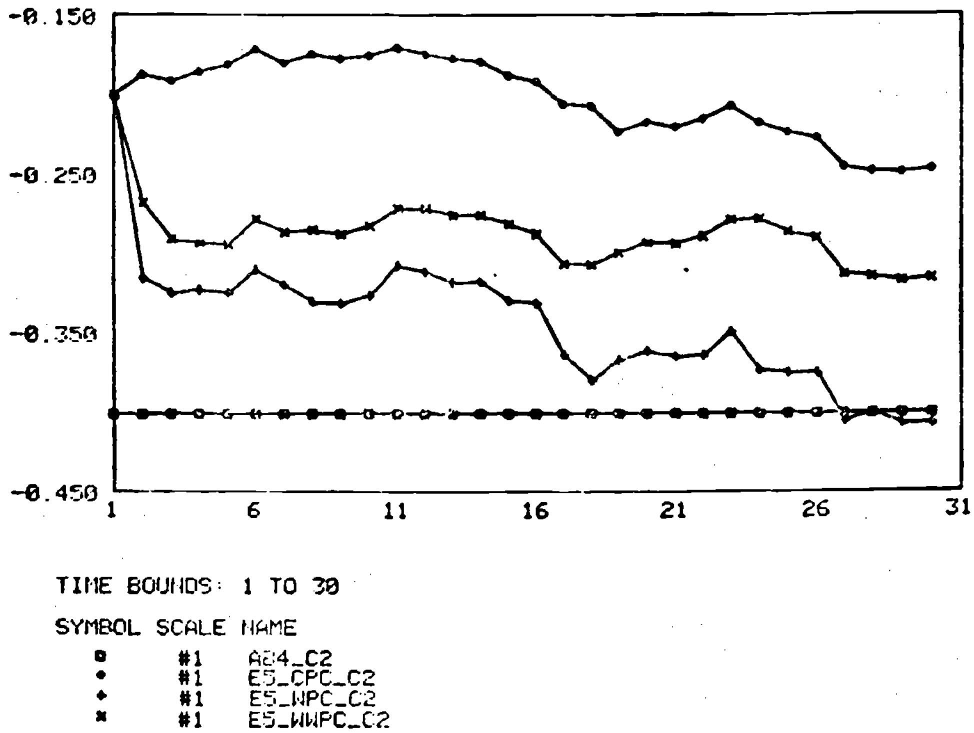


TIME BOUNDS: 1 TO 30

SYMBOL SCALE NAME

■ #1 AB4_C1
• #1 E5_WMPC_C1
♦ #1 E5_S1PC_C1

Figure 11. Estimates of a_t in E5.



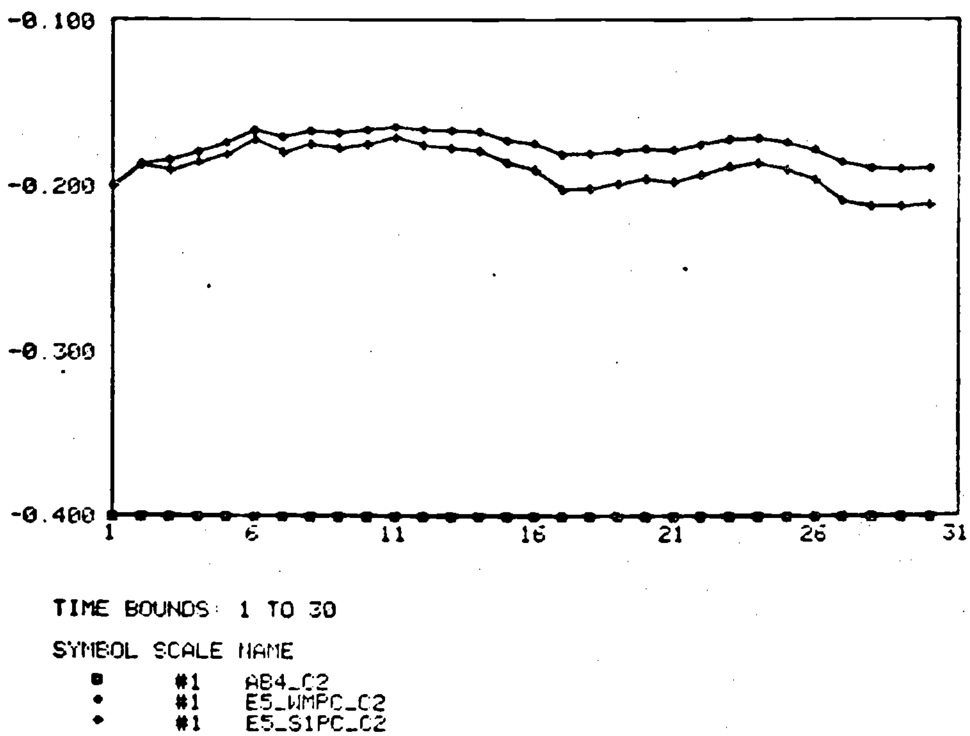


Figure 12. Estimates of b_t in E5.

9. SUMMARY AND CONCLUSIONS.

In this paper we have examined the problem of controlling a system with parameters varying in a fashion unknown to the controller. We have surveyed all methods available for the solution of such problems and we have extended some to fit our framework. We have also suggested and analyzed two methods for the first time. One is a non-dual one (S1) and the other is a dual one (see section 7.3). We have also presented a numerical comparison of the non-dual methods, in which S1 was found, along with enforced separation, superior to other non-dual methods that have been suggested elsewhere.

A major problem with all the methods is that a-priori there is complete ignorance about the evolution of the parameters. From figures 5 and 6 it was seen that if the parameter variation happens to be of the same form as the one assumed, then these parameters are estimated satisfactorily. Otherwise, we do not have large hope of identifying them. This raises the whole issue of robust estimation for some particular kind of parameter variation, it is not clear whether it will give good results if the parameters evolve according to a different structure. The ultimate goal, of course, is to optimize the criterion. The interaction between identification and control might be somewhat understood in the case of constant but unknown parameters, but it is not at all clear in the case of time varying parameters. There is still a lot of research to be done in this area beginning with more extensive comparisons of the dual and the non-dual methods, extensions to higher order systems, and examination of the interaction between identification and control.

APPENDIX A

SOLUTION OF THE ADAPTIVE COVARIANCE CONTROL PROBLEM

In this appendix we present the solution to the problem posed in section 7.2. The solution procedure follows the analysis of McRae (1972). The problem is the following.

Find $u_k, u_{k+1}, \ldots, u_{N-1}$ where

$$V^*(y^k) = \min_{u_k, u_{k+1}, \ldots u_{N-1}} E\left\{ \sum_{i=k}^{N-1} (y_{i+1}^2 + r u_i^2)/k-1 \right\} \qquad (A.1)$$

subject to

$$y_{j+1} = z_j p_j + \varepsilon_j \qquad j \geq k \qquad (A.2)$$

ε_j independent zero mean white noise with covariance σ_ε^2

$$p(p_j/j) = N(p_{j/j}, M_{j/j}) \qquad (A.3)$$

$$p_{j/j} = p_{j/j-1} = p_{j-1/j-1} = \cdots = p_{k/k-1} \qquad (A.4)$$

$$M_{j/j}^{-1} = M_{j/j-1}^{-1} + E\left[\frac{1}{\sigma_\varepsilon^2} z'_j z_j /k-1\right] \qquad j \geq k \qquad (A.5)$$

$$M_{j/j-1} = M_{j-1/j-1} + R \qquad j \geq k \qquad (A.6)$$

$$z_j = [y_j, u_j] \qquad j \geq k \qquad (A.7)$$

We define a set of N-k+1 matrix Lagrange multipliers L_j $k-1 \leq j \leq N-1$ where L_j are all symmetric 2x2 matrices. We now form the following Hamiltonian quantity.

$$H(y^k) = E\{\sum_{i=k}^{N-1} (y_{i+1}^2 + ru_i^2)/k-1\} +$$

$$\sum_{i=k}^{N-1} tr\{ L_i \left[M^{-1}_{i/i} - (M_{i-1/i-1} + R)^{-1} - E(\frac{1}{\sigma_\varepsilon^2} z'_i z_i /k-1)\right] \} =$$

$$\sum_{i=k}^{N-1} \{ E\left[y_{i+1}^2 + ru_i^2 - \frac{1}{\sigma_\varepsilon^2} tr(L_i z'_i z_i)/k-1\right] + \Delta_i \} +$$

$$tr\left[L_{N-1}M^{-1}_{N-1/N-1} - L_{k-1}M^{-1}_{k-1/k-1}\right] \quad (A.8)$$

where

$$\Delta_i = tr\left[L_{i-1}M^{-1}_{i-1/i-1} - L_i (M_{i-1/i-1} + R)^{-1}\right] \quad (A.9)$$

We shall apply stochastic dynamic programming to the augmented criterion (A.8). We shall be careful, however, to simultaneously impose the constraints

$$\frac{\partial H(y^k)}{\partial M^{-1}_{j/j}} = 0 \qquad k \leq j \leq N-1 \quad (A.10)$$

$$\frac{\partial H(y^k)}{\partial L_j} = 0 \qquad k \leq j \leq N-1 \quad (A.11)$$

(A.10) and (A.11) correspond to the state-costate equations for $M_{j/j}$.

The dynamic programming recursion can now be written as follows

$$H^*(y^j) = \min_{u_j} \{ E\left[W_j + H^*(y^{j+1})/j-1\right] + \Delta_j\} \quad (A.12)$$

for

$k \leq j \leq N-1$

with

$$H^*(y^N) = tr\left[L_{N-1}M^{-1}_{N-1/N-1} - L_{k-1}M^{-1}_{k-1/k-1}\right] \quad (A.13)$$

$$W_j = y^2_{j+1} + ru^2_j - \frac{1}{\sigma^2_\varepsilon} tr(L_j z'_j z_j) \quad (A.14)$$

$$\Delta_j = tr\left[L_{j-1}M^{-1}_{j-1/j-1} - L_j(M_{j-1/j-1} + R)^{-1}\right] \quad (A.15)$$

$$E\left[./j-1\right] = E\left[./y^j, u^{j-1}\right] \quad (A.16)$$

The interesting thing about this arrangement is that we shall be able to satisfy (A.10) recursively as we proceed backwards.

At time N we have the cost $H^*(y^N)$ given in (A.13). We can differentiate it with respect to $M^{-1}_{N-1/N-1}$ since this quantity will appear only in $H^*(y^N)$. Using (B.4) of appendix B, we have

$$\frac{\partial H^*(y^k)}{\partial M^{-1}_{N-1/N-1}} = \frac{\partial H^*(y^N)}{\partial M^{-1}_{N-1/N-1}} = L_{N-1} + L'_{N-1} - DIAG(L_{N-1}) = 0 \quad (A.17)$$

since the L_j are symmetric (A.17) is equivalent to

$$L_{N-1} = 0 \quad (A.18)$$

$$H^*(y^N) = - trL_{k-1}M^{-1}_{k-1/k-1} \quad (A.19)$$

At time N-1 the recursion (A.12) becomes

$$H^*(y^{N-1}) = \min_{u_{N-1}} \{ E\left[y^2_N + ru^2_{N-1} - \frac{1}{\sigma^2_\varepsilon} tr(L_{N-1}z'_{N-1}z_{N-1}) + H^*(y^N)/N-2\right] + tr\left[L_{N-2}M^{-1}_{N-2/N-2} - L_{N-1}(M_{N-2/N-2} + R)^{-1}\right]\} \quad (A.20)$$

Only the first three terms in (A.20) involve u_{N-1}. We partition the matrices L_j and M_j as follows

$$L_j = \begin{bmatrix} L_j^a & L_j^{ab} \\ L_j^{ab} & L_j^b \end{bmatrix} \tag{A.21}$$

$$M_{j/j} = \begin{bmatrix} M_{j/j}^a & M_{j/j}^{ab} \\ M_{j/j}^{ab} & M_{j/j}^b \end{bmatrix} \tag{A.22}$$

We now expand (A.20)

$$H^*(y^{N-1}) = \min_{u_{N-1}} \{y_{N-1}^2(a_{N-1/N-2}^2 + M_{N-1/N-2}^a) + 2y_{N-1}u_{N-1}(a_{N-1/N-2}b_{N-1/N-2} + M_{N-1/N-2}^{ab}) + u_{N-1}^2(r + b^2_{N-1/N-2} + M_{N-1/N-2}^b) + \sigma_\varepsilon^2 - \frac{1}{\sigma_\varepsilon^2}(L_{N-1}^a y_{N-1}^2 + 2L^{ab}_{N-1}y_{N-1}u_{N-1} + L_{N-1}^b u_{N-1}^2) + H^*(y^N) + tr\left[L_{N-2}M^{-1}_{N-2/N-2} - L_{N-1}(M_{N-2/N-2} + R)^{-1}\right]\} \tag{A.23}$$

By differentiation we find that the minimizing u_{N-1} is

$$u_{N-1}^* = -G^{-1}_{N-1}F_{N-1}y_{N-1} \tag{A.24}$$

where

$$G_{N-1} = r + b_{N-1/N-2}^2 + M_{N-1/N-2}^b - \frac{1}{\sigma_\varepsilon^2} L_{N-1}^b \tag{A.25}$$

$$F_{N-1} = a_{N-1/N-2}b_{N-1/N-2} + M^{ab}_{N-1/N-2} - \frac{1}{\sigma_\varepsilon^2} L^{ab}_{N-1} \qquad (A.26)$$

We now write $H(y^k)$ in a form that will help the differentiation with respect to $M^{-1}_{N-2/N-2}$ dictated by (A.10)

$$H(y^k) = E\left[y_N^2/k-1\right] + tr\left[L_{N-2}M^{-1}_{N-2/N-2} - L_{N-1}\{M_{N-2/N-2} + R)^{-1}\right] +$$

(terms not involving $M_{N-2/N-2}$) =

$$trE(z_{N-1}P_{N-1}P'_{N-1}z'_{N-1}/k-1) + tr\left[L_{N-2}M^{-1}_{N-2/N-2} - L_{N-1}(M_{N-2/N-2} + R)^{-1}\right]$$

$$+ \text{(terms without } M_{N-2/N-2}) = tr\left[(P_{N-1/N-2}P'_{N-1/N-2} + M_{N-2/N-2} + R) \cdot\right.$$

$$\left.\cdot E(z'_{N-1}z_{N-1}/k-1)\right] + tr\left[L_{N-2}M^{-1}_{N-2/N-2} - L_{N-1}(M_{N-2/N-2} + R)^{-1}\right] +$$

$$\text{(terms without } M_{N-2/N-2}) \qquad (A.27)$$

With the help of (B.5) and (B.6) we have

$$\frac{\partial H(y^k)}{\partial M^{-1}_{N-2/N-2}} = 0 = -2M_{N-2/N-2}E(z'_{N-1}z_{N-1}/k-1)M_{N-2/N-2} +$$

$$DIAG\left[M_{N-2/N-2}E(z'_{N-1}z_{N-1}/k-1)M_{N-2/N-2}\right] - 2L_{N-2} + DIAG(L_{N-2}) +$$

$$2(I + RM^{-1}_{N-2/N-2})^{-1}L_{N-1}(I + M^{-1}_{N-2/N-2}R)^{-1} -$$

$$\text{DIAG}\left[(I + RM^{-1}_{N-2/N-2})^{-1}L_{N-1}(I + M^{-1}_{N-2/N-2}R)^{-1}\right] \qquad (A.28)$$

Since all the matrices are symmetric, (A.28) is equivalent to the following

$$L_{N-2} = (I + RM^{-1}_{N-2/N-2})^{-1}L_{N-1}(I + M_{N-2/N-2}R)^{-1} - M_{N-2/N-2}E\left[z'_{N-1}z_{N-1}/k-1\right] \cdot$$

$$\cdot \quad M_{N-2/N-2} \qquad (A.29)$$

The cost $H^*(y^{N-1})$ becomes

$$H^*(y^{N-1}) = y^2_{N-1}K_{N-1} + \sigma^2_\varepsilon + H^*(y^N) + tr\left[L_{N-2}M^{-1}_{N-2/N-2} - L_{N-1}(M_{N-2/N-2} + R)^{-1}\right] \quad (A.30)$$

where

$$K_{N-1} = a^2_{N-1/N-2} + M^a_{N-1/N-2} - \frac{1}{\sigma^2_\varepsilon} L^a_{N-1} - G^{-1}_{N-1}F^2_{N-1} \qquad (A.31)$$

So $H^*(y^{N-1})$ is a quadratic in y_{N-1} and the recursion can continue. Notice that the nonlinear dependence of K_{N-1} on $u_{N-2},\ldots.u_k$ has dissappeared with the introduction of the multiplier matrices. It is easy now to write the expressions for u^*_j.

$$u^*_j = -G^{-1}_j F_j y_j \qquad (A.32)$$

where

$$G_j = r + (1 + K_{j+1})(b^2_{j/j-1} + M^b_{j/j-1}) - \frac{1}{\sigma^2_\varepsilon} L^b_j \qquad (A.33)$$

$$F_j = (1 + K_{j+1})(a_{j/j-1}b_{j/j-1} + M^{ab}_{j/j-1}) - \frac{1}{\sigma^2_\varepsilon} L^{ab}_j \qquad (A.34)$$

$$K_j = (1 + K_{j+1})(a^2_{j/j-1} + M^a_{j/j-1}) - \frac{1}{\sigma^2_\varepsilon} L^a_j - G^{-1}_j F^2_j \qquad (A.35)$$

$$L_j = (I + RM^{-1}_{j/j})L_{j+1}(I + M^{-1}_{j/j}R^{-1} - M_{j/j}E(z'_{j+1}z_{j+1}/k-1)M_{j/j} \qquad (A.36)$$

The initial conditions are

$$K_N = 0 \;,\; L_{N-1} = 0 \qquad (A.37)$$

Along with (A.4), (A.5) and (A.6) the above equations define a complicated two-point-boundary-value (TPBV) problem. In order to define the problem completely we need a way to evaluate

$E(z'_j z_j/k-1)$ for all $k \le j \le N-1$

We now provide such a recursion.

$$E(z'_j z_j/k-1) = \begin{bmatrix} 1 & -G^{-1}_j F_j \\ -G^{-1}_j F_j & G^{-2}_j F^2_j \end{bmatrix} E(y^2_j/k-1) \equiv P_j E(y^2_j/k-1) \qquad (A.38)$$

$$E(y^2_j/k-1) = \sigma^2_\varepsilon + E(z_{j-1}P_{j-1}P'_{j-1}z'_{j-1}/k-1) =$$

$$= \sigma_\varepsilon^2 + \text{tr}\left[P_{j-1}(P_{j-1/j-2}P'_{j-1/j-2} + M_{j-1/j-2})\right] \cdot E(y^2_{j-1}/k-1) \quad \text{(A.39)}$$

Since $E(y_k^2/k-1) = y_k^2$ (A,39) is a well defined recursion. The TPBV problem is now complete.

APPENDIX B

SOME USEFUL MATRIX DERIVATIVES

In this appendix we develop certain matrix derivatives that are useful in the proofs of appendix A. Many formulas for matrix derivatives have been reported by Athans and Schweppe (1965), and Athans (1967). However, those derivatives were applicable only to matrices whose elements are independent. Here we derive some formulas for symmetric matrices.

Define the operator DIAG which operates on a square matrix A and creates the following matrix

$$\text{DIAG}\,(A) = \begin{bmatrix} a_{11} & 0 & & \\ 0 & a_{22} & & \underline{0} \\ \underline{0} & & & a_{nn} \end{bmatrix} \tag{B.1}$$

Let X be a nxn matrix and let f(X) denote a scalar valued function of the n^2 elements of X. Then the matrix derivative of f is defined by

$$\frac{\partial}{\partial X} f(X) = \left\{ \frac{\partial f(X)}{\partial X_{ij}} \right\} \tag{B.2}$$

so the matrix derivative of f is a matrix. We now state the following theorem.

<u>Theorem</u>. Let X,B be symmetric nxn matrices. Then the following equalities are true

$$\frac{\partial}{\partial X} trX = I \tag{B.3}$$

$$\frac{\partial}{\partial X} trAX = A + A' - DIAG(A) \tag{B.4}$$

$$\frac{\partial}{\partial X} trA(X+B)^{-1} = -(X+B)^{-1}A(X+B)^{-1} - (X+B)^{-1}A'(X+B)^{-1} +$$

$$DIAG\left[(X+B)^{-1}A(X+B)^{-1}\right] \tag{B.5}$$

$$\frac{\partial}{\partial X} trA\left[X^{-1}+B\right]^{-1} = (I+BX)^{-1}A(I+XB)^{-1} + (I+BX)^{-1}A'(I+XB)^{-1} -$$

$$DIAG\left[(I+BX)^{-1}A(I+XB)^{-1}\right] \tag{B.6}$$

Before we proceed with the proofs we state for comparison the corresponding formulas for matrices whose elements are independent

$$\frac{\partial}{\partial X} trX = I \tag{B.7}$$

$$\frac{\partial}{\partial X} trAX = A' \tag{B.8}$$

$$\frac{\partial}{\partial X} trA(X+B)^{-1} = -\left[(X+B)^{-1}A(X+B)^{-1}\right]' \tag{B.9}$$

$$\frac{\partial}{\partial X} trA(X^{-1}+B)^{-1} = \left[(I+BX)^{-1}A(I+XB)^{-1}\right]' \tag{B.10}$$

<u>Proof</u>. (B.3) is trivial and we omit its proof.

(B.4):

$$\frac{\partial}{\partial X_{ij}} \operatorname{tr} AX = \operatorname{tr} A \frac{\partial X}{\partial X_{ij}} = \operatorname{tr} A \begin{bmatrix} 0 & & \\ & 0 & 1_{ij} \\ 1_{ji} & & 0 \end{bmatrix} =$$

$$a_{ij} + a_{ji} \qquad \text{Q.E.D.}$$

(B.5):

$$\frac{\partial}{\partial X_{ij}} \operatorname{tr} A(X + B)^{-1} = \operatorname{tr} A \frac{\partial}{\partial X_{ij}} (X + B)^{-1} = -\operatorname{tr} A(X + B)^{-1} \frac{\partial X}{\partial X_{ij}} (X + B)^{-1} =$$

$$-\operatorname{tr} A' \{X^{jk}X^{li} + X^{ik}X^{lj}\} \qquad (\{k,l\} = 1.2.,..,n.)$$

where

$$X^{mp} \equiv \left[(X + B)^{-1}\right]_{mp}$$

$$-\operatorname{tr} A' \{X^{jk}X^{li} + X^{ik}X^{lj}\} =$$

$$-\sum_{l=1}^{n} \sum_{k=1}^{n} a_{lk}(X^{jl}X^{ki} + X^{il}X^{kj}) =$$

$$-\left[(X + B)^{-1}A(X + B)^{-1} + (X + B)^{-1} A'(X + B)^{-1}\right]_{ij} \qquad \text{Q.E.D.}$$

(B.6): $$\frac{\partial}{\partial X_{ij}} \operatorname{tr}A(X^{-1} + B)^{-1} = \operatorname{tr}A \frac{\partial (X^{-1} + B)^{-1}}{\partial X_{ij}} =$$

$$-\operatorname{tr}A(X^{-1} + B)^{-1} \frac{\partial X^{-1}}{\partial X_{ij}} (X^{-1} + B)^{-1} =$$

$$\operatorname{tr}A(X^{-1} + B)^{-1} X \frac{\partial X}{\partial X_{ij}} X^{-1}(X^{-1} + B)^{-1} =$$

$$\operatorname{tr}A(I + XB)^{-1} \frac{\partial X}{\partial X_{ij}} (I + BX)^{-1}$$ and the analysis of

the proof of (B.5) carries over.

APPENDIX C

COMPUTATION OF THE TWO-STEP ADAPTIVE CONTROL

In this appendix we carry out the calculations called for in section 7.3. The problem is

$$V^*(y^k) = \min_{u_k} E\left[y_{k+1}^2 + ru_k^2 + Hy_{k+1}^2 + F/k-1\right] \tag{C.1}$$

We substitute (121) and (122) into (118) keeping in mind the assumption (124). We obtain

$$H = \frac{r}{r + b_{k+1}^2} \{a_{k/k-1}^2 + M^{a^2}_{k/k-1} y_k^2 (y_k^2 M^a_{k/k-1} + \sigma_\varepsilon^2)^{-2} \left[(a_k - a_{k/k-1})y_k + \varepsilon_k\right]^2 + 2a_{k/k-1} M^a_{k/k-1} y_k (y_k^2 M^a_{k/k-1} + \sigma_\varepsilon^2)^{-1} \left[(a_k - a_{k/k-1})y_k + \varepsilon_k\right]\} + + M^a_{k+1/k} \tag{C.2}$$

We notice from (122) that $M^a_{k+1/k}$ does not depend on u_k and is a function of (y^k, u^{k-1}) so we will not expand it further. To facilitate the notation we shall define the quantity

$$X_k \equiv M^a_{k/k-1} y_k (y_k^2 M^a_{k/k-1} + \sigma_\varepsilon^2)^{-1} \tag{C.3}$$

X_k is a function of (y^k, u^{k-1}).

$V(y^k)$ now becomes

$$V(y^k) = E\{ru_k^2 + F + (a_k^2 y_k^2 + b_k^2 u_k^2 + \varepsilon_k^2 + 2a_k b_k y_k u_k +$$

$$2a_k y_k \varepsilon_k + 2b_k u_k \varepsilon_k)\{1 + M^a_{k+1/k} +$$

$$\frac{r}{r + b^2_{k+1}} \left[a^2_{k/k-1} + X_k^2 \left[(a_k - a_{k/k-1}) y_k + \varepsilon_k \right]^2 + 2a_{k/k-1} X_k \right.$$

$$\left. \left[(a_k - a_{k/k-1}) y_k + \varepsilon_k \right] \right] \} / y^k, u^{k-1} \} =$$

$$ru^2_k + F + y_k^2 \{ (1 + M^a_{k+1/k})(a^2_{k/k-1} + M^a_{k/k-1}) +$$

$$\frac{r}{r + b^2_{k+1}} \left[a^2_{k/k-1} (a^2_{k/k-1} + M^a_{k/k-1}) + X_k^2 \left[y_k^2 E \left[a_k^2 (a_k - a_{k/k-1})^2 / k-1 \right] + \right. \right.$$

$$\left. \left. \sigma_\varepsilon^2 (a^2_{k/k-1} + M^a_{k/k-1}) \right] + 2a_{k/k-1} X_k y_k E \left[a_k (a_k - a_{k/k-1}) / k-1 \right] \right] \} +$$

$$u_k^2 \{ (1 + M^a_{k+1/k})(b^2_{k/k-1} + M^b_{k/k-1}) + \frac{r}{r + b^2_{k+1}} \left[a^2{}_{k/k-1} (b^2_{k/k-1} + \right.$$

$$\left. M^b_{k/k-1}) + X_k^2 \left[E \left[b_k^2 (a_k - a_{k/k-1})^2 / k-1 \right] y_k^2 + \sigma_\varepsilon^2 (b^2_{k/k-1} + M^b_{k/k-1}) \right] \right] \} +$$

$$(1 + M^a_{k+1/k}) \sigma_\varepsilon^2 + \frac{r}{r + b^2_{k+1}} \left[a^2_{k/k-1} \sigma_\varepsilon^2 + X_k^2 \left[M^a_{k/k-1} y_k^2 \sigma_\varepsilon^2 + \bar{\varepsilon}_k^4 \right] \right] +$$

$$2y_k u_k \{ (1 + M^a_{k+1/k}) E(a_k b_k/k-1) + \frac{r}{r + b^2_{k+1}} \Big[a^2_{k/k-1} E(a_k b_k/k-1) +$$

$$X^2_k \Big[y^2_k E \big[a_k b_k (a_k - a_{k/k-1})^2/k-1 \big] + \sigma^2_\varepsilon E(a_k b_k/k-1) \Big] +$$

$$2a_{k/k-1} X_k y_k E \big[a_k b_k (a_k - a_{k/k-1})/k-1 \big] \} +$$

$$2y_k \left(\frac{r}{r + b^2_{k+1}} \right) \Big[2X^2_k E \big[a_k (a_k - a_{k/k-1})/k-1 \big] \sigma^2_\varepsilon y_k + 2a^2_{k/k-1} X_k \sigma^2_\varepsilon \Big] +$$

$$2u_k \left(\frac{r}{r + b^2_{k+1}} \right) \Big[2X^2_k y_k E \big[b_k (a_k - a_{k/k-1})/k-1 \big] \sigma^2_\varepsilon + 2a_{k/k-1} b_{k/k-1} X_k \sigma^2_\varepsilon \Big] \quad (C.4)$$

Equation (C.4) is a quadratic in u_k so its minimization is straightforward. We find

$$u^*_k = -D^{-1} f_k \qquad (C.5)$$

where (remembering that b_{k+1} was assumed equal to $b_{k/k-1}$)

$$D_k = r + (1 + M^a_{k+1/k})(b^2_{k/k-1} + M^b_{k/k-1}) + \frac{r}{r + b^2_{k/k-1}} \{ a^2_{k/k-1} (b^2_{k/k-1}$$

$$+ M^b_{k/k-1}) + X^2_k \Big[E \big[b^2_k (a_k - a_{k/k-1})^2/k-1 \big] y^2_k + \sigma^2_\varepsilon (b^2_{k/k-1} + M^b_{k/k-1}) \Big] \} \quad (C.6)$$

$$f_k = y_k \{(1 + M^a_{k+1/k})E(a_k b_k/k-1) + \frac{r}{r + b^2_{k/k-1}} \left[a^2_{k/k-1}E(a_k b_k/k-1) + \right.$$

$$X^2_k \left[y^2_k E\left[a_k b_k (a_k \; a_{k/k-1})^2/k-1\right] + \sigma^2_\varepsilon E(a_k b_k/k-1)\right] +$$

$$2a_{k/k-1}X_k y_k E\left[a_k b_k (a_k - a_{k/k-1})/k-1\right] \} +$$

$$\frac{r}{r + b^2_{k/k-1}} \left[2X^2_k \; y_k E\left[b_k (a_k - a_{k/k-1})/k-1\right]\sigma^2_\varepsilon + 2a_{k/k-1}b_{k/k-1}X_k \sigma^2_\varepsilon\right] \quad (C.7)$$

The expectations appearing in D_k and f_k are straightforward to compute from the joint gaussian density of a_k and b_k given (y^k, u^{k-1}).

REFERENCES

1. Aoki, M. (1967), "Optimization of Stochastic Systems," Academic Press, 1967.

2. Åström, K. J. (1970), Introduction to Stochastic Control Theory, Academic Press, 1970.

3. Athans, M. (1967), "The Matrix Minimum Principle," Information and Control, Vol. II, Nos. 5 and 6, November-December 1967.

4. Athans, M. and F. C. Schweppe (1965), "Gradient Matrices and Matrix Calculations," MIT Lincoln Lab. Technical Note 1965-53, Lexington, Massachusetts.

5. Bar-Shalom, Y. and Sivan, R. (1969), "On the Optimal Control of Discrete-Time Linear Systems with Random Parameters," IEEE Transactions on Automatic Control, Vol. AC-14, No. 1, February 1969.

6. Chow, G. C. (1973), Private Communication.

7. Cooley, T. F. (1971), "Estimation in the Presence of Sequential Parameter Variation," Ph. D. Thesis, Department of Economics, University of Pennsylvania, 1971.

8. Curry, R. E. (1969), "A New Algorithm for Suboptimal Stochastic Control," IEEE Transactions on Automatic Control, Vol. AC-14, No. 5, October 1969.

9. Curry, R. E. (1970), Estimation and Control with Quantized Measurements, MIT Press 1970.

10. Early, R. H. and B. M. Early (1973), "On the Relative Performance of the Optimal Control System with M Measurements," IEEE Transactions on Automatic Control, Vol. AC-17, No. 4, August, 1973.

11. Fel'dbaum, A. A. (1960a), "Dual Control Theory 1," Automation and Remote Control, Vol. 21, No. 11, September 1960.

12. Fel'dbaum, A. A. (1960b), "Dual Control Theory 2," Automation and Remote Control, Vol. 21, No. 11, November 1960.

13. Fel'dbaum, A. A. (1961a), "Dual Control Theory 3," Automation and Remote Control, Vol. 22, No. 1, January 1961.

14. Fel'dbaum, A. A. (1961b), "Dual Control Theory 4," Automation and Remote Control, Vol. 22, No. 2, February 1961.

15. Fel'dbaum, A. A. (1966), "Optimal Control Systems," Academic Press, New York, 1966.

16. IEEE Transactions on Automatic Control, "Special Issue on Linear-Quadratic-Gaussian Problem," Vol. AC-16, December 1971.

17. Ku, R. and M. Athans (1973), "On the Adaptive Control of Linear Systems Using the Open-Loop-Feedback-Optimal Approach," IEEE Transactions on Automatic Control, Vol. AC-18, No. 5, October 1973.

18. McRae, E. C. (1972), "Linear Decision with Experimentation," Annals of Economic and Social Measurement, October 1972.

19. Popović, B. (1972), "Analytic Solutions to Some Stochastic Adaptive, and Dual Optimal Control Problems for Economic Decision Making," Ph. D. Thesis, University of Illinois, Urbana-Champaign.

20. Prescott, E. C. (1967), "Adaptive Decision Rules for Macro Economic Planning," Ph. D. Thesis, Graduate School of Industrial Administration, Carnegie-Mellon University, 1967.

21. Rausser, G. C. and J. W. Freebairn (1973), "Comparison of Approximate Adaptive Control Solution to U. S. Beef Trade Policy," Presented at the second NBER Conference on Stochastic Control and Economics, University of Chicago, June 7-9, 1973.

22. Rosenberg, B. (1968), "Varying-Parameter Estimation," Ph. D. Thesis, Department of Economics, Harvard University 1968.

23. Sarris, A. H. (1973), "A Bayesian Approach to Estimation of Time-Varying Regression Coefficients," Annals of Economic and Social Measurement, October 1973.

24. Theil, H. (1971), <u>Principles of Econometrics</u>. John Wiley, 1971.

25. Tse, E. and M. Athans (1972), "Adaptive Stochastic Control for a Class of Linear Systems," IEEE Transaction on Automatic Control, Vol. AC-17, No. 1, February 1972.

26. Tse, E., Y. Bar-Shalom and L. Meier, III (1973a), "Wide Sense Adaptive Dual Control for Nonlinear Stochastic Systems," IEEE Transactions on Automatic Control, Vol. AC-18, No. 2, April 1973.

27. Tse, E. and Y. Bar-Shalom (1973b), "An Activity Adaptive Control for Linear Systems with Random Parameters via the Dual Control Approach," IEEE Transaction on Automatic Control, Vol. AC-18 No. 2, April 1973.

28. Wieslander, J. and B. Wittenmark (1971), "An Approach to Adaptive Control Using Real Time Identification," Automatica, Vol. 7, May 1971.

29. Wouters, W. (1973), "Adaptive Minimum Variance Control for Linear Discrete Time Systems," Proceedings of the 1972 IEEE Conference on Decision and Control, December 13-15, 1972.

30. Yoshida, Y. and Nakamura, K., (1973), "Learning Dual Control with Complete State Information," Research Reports of Automatic Control Laboratory Faculty of Engineering, Nagoya University, Japan, Vol. 20, June 1973.